

Computers in the office

Office automation

An office is any place where business is carried out. It's equipment for many years has been no more than a telephone, a typewriter and a few filing cabinets. It is no wonder, then, that the automated office is composed of electronic versions of these basic components.

Office automation was a term coined a number of years ago to define what it was that computers were trying to do to the typical office environment. The phrase

brings to mind the picture of offices as small factories; in some respects this is true, the product being **information** which is manipulated as on a production line.

Office automation is more than just removing the boredom from routine work. It is an *approach*, a new way of handling information. An automated office consists of many components linked together in such a way that once information is entered into the system, it can be processed and channelled from place to place with little human intervention.

Below: the Rainbow 100 personal computer.
(Photo: Digital Equipment Company Ltd).





All information that comes into the office is a candidate for improved efficiency of handling through automation, for example, information regarding job applications. Once a candidate's name and address has been typed into the system, that data can then be called up every time a letter needs sending or when information needs to be printed out.

Once a successful candidate has been chosen, the data already stored about him or her can then form part of the personnel department's files. The advertisement that was initially prepared by personnel may even form part of the new employee's job description.

At no stage in this process has a computer been selecting potential employees by qualification (although this is feasible). It is important that the computers themselves be easy to use and as trouble

free as possible. Systems designers, therefore, must always take the user into account when designing hardware and software. The system would not be functioning efficiently if it took longer to operate than the manual system it had replaced!

What is an office?

The needs of modern commercial and governmental practice are changing. In order to capitalise on opportunity and respond quickly to changing circumstances, information needs to be sent and received as quickly as possible, regardless of location. This has led to the development of *portable* communications terminals. Now, the office is wherever you happen to be with your portable computer – a telephone modem link connects you to the central information store and messages

Above: ACT Apricot business computer. (Photo: ACT).

and updated information can be transmitted at any time. This is particularly useful when business is being conducted across time zones.

Portable terminals

Portable communications terminals range from the personal computers often called **executive workstations**, to full size microcomputers that really need wheels to be truly portable. (A better name for these might be transportable computers.)

Executive workstations are small VDUs, usually about 9 inches in diameter (compared with the normal 12 or 15 inch VDUs). They are attached to electronic telephones that can be programmed to dial lists of numbers automatically on demand. Some also possess an RS 232C serial plug port, enabling data from other computers to be displayed on screen.

Portable workstations now often use liquid crystal displays rather than CRTs. The advantages are: the display is physically smaller for a given display area (it comes in a flatpack); less electronics is needed to drive the display; less power and no high tension supply is required. Many are able to link up with a small heat sensitive printer so that the executive can obtain hard copy.

More sophisticated, yet still primitive, portable terminals incorporate a telephone handset with a full sized keyboard. Such models generally offer two keyboards: one full size for word processing; the other somewhat smaller for the business person who only wants to press the odd key ('y' and 'no', for example) to access and manipulate the data in a database.

The most sophisticated terminals at the moment handle electronic mail and the receipt or transmission of telex messages. Some can also manipulate data from Prestel and other viewdata systems.

Electronic mail – an introduction

In an office, where, say, 40 engineers each have a computer terminal linked to a mainframe host, an electronic mail system can take care of all their internal memos, design queries, information searches and many other tasks. Conversational mode is possible but it means a lot of typing – it is still quicker to use the phone.

The most common use for electronic

mail is outside of one building and across international boundaries – this is mainly because the rates for data transmission are cheaper (and the connection is shorter) than for voice.

As this automated design office develops and grows, the company creates probably its most valuable resource: data. Company contracts, invaluable programs, design specifications, customer specifications, accounts – all this needs to be protected in the same way that money would be safeguarded.

Some companies take back-up copies of their databases which are then stored in a bank. Another method of securing data is to use some form of data encryption to encode the data. In this way, only authorised personnel can read and change it. (See *Computer Science 13*.)

Making computers accessible to so many people means that there is a continual demand for data storage space – efficient data management systems are therefore necessary.

The paperless office?

Paper is the usual medium for storing information and moving it around. However, in comparison with computerised data it is expensive, and difficult and time consuming to extract useful information from.

However, the automated office is not really concerned with eliminating paper from the workplace. What it does promise, though, is the ability to manage information more quickly and more productively.

Office computer systems

It is the ability of computers to store and keep track of large quantities of data which can be accessed quickly that makes them invaluable for business. And indeed, since they were first manufactured in the 1940s and '50s, computers have always been used in offices.

A mainframe computer stores data on reels of magnetic tape or on stacks of magnetic disks. For the last twenty years or so they have been carrying out the routine task of sending out rates demands and processing weekly payrolls for the large companies that could afford them.

However, they are expensive to run, they require air conditioning to keep them



Above: Fortune 32:16 professional computer. (Photo: TIS Computing Ltd).

cool and specially trained operators are needed to program and maintain them. Because of this, mainframes were housed in small 'back' rooms with the result that few people were able to get access to them whenever they needed to.

This difficulty prompted a flood of mini and now microcomputer manufacturers to deluge the office marketplace with smaller, less powerful and cheaper computers. Now, anyone that wanted a machine could have one. There are now over 250 computer manufacturers, and more than 10,000 software applications packages are available to run on their machines.

Individual mini or microcomputers though, have the disadvantage in that they *cannot* communicate – information can only be transferred from machine to

machine via their removable storage media, floppy disks for example.

The real key to the office of the future is **communications**. By connecting all these individual personal computers, micros, minis and executive workstations into a **data communications network**, users will be able to 'talk' to each other. They will also be able to communicate with mainframes, downloading sections of a database and transferring files.

To summarise, then, the main components of the automated office will be: data capture (how information gets into the system in the first place); data storage (how and where the office data will be stored); and data communications (how that data is moved from place to place and the equipment that carries it out).

Data capture

Although data files are in many ways the centre of any office system, there must always be some method of putting data into the system in the first place. The task of collecting data and recording it in a form suitable for a computer is known as data capture. Data input is the operation that encodes information and stores it in the computer, although the term is often used to mean data capture.

The organisation of data capture must take into consideration the volume, frequency and type of data, among other things. There are also important human considerations, such as the skills of existing staff and the acceptable levels of accuracy and update. We will describe the characteristics of various input devices to make it clear just what facilities are available.

On-line terminals

Terminals used to be teletype machines – keyboard and printer input/output devices – which interfaced directly with computers, conveying human information into the machine.

When cathode ray tubes became cheap and easy to produce, however, they were seen as a possible replacement for the teletype terminal. The CRT or VDU soon replaced the teletype and it became known as the glass teletype, or dumb VDU.

The VDU uses the same communications language or **protocol** as the teletype and, in effect, the two are simply the media by which the operator enters commands and data. The replacement of the teletype by the CRT was so complete that the word 'terminal' really means VDU now.

Dumb VDUs are limited in application for the office of the future, because no editing can be carried out locally on them. As soon as a command or some data is typed into the machine, it is immediately sent off to the host for processing or storage. At the other end of the terminal market are the smart or **intelligent terminals**. Because of the range of functions that these now provide, they resemble personal computers; sometimes called workstations.

These intelligent terminals may be used as conventional dumb VDUs to send and receive messages, but they can also

carry out processing locally, running jobs under the CP/M operating system, for instance.

Key-to-store machines

The use of intelligent, as opposed to dumb, on-line terminals as interactive input devices can still interrupt the work of the central processor. If used extensively they may seriously degrade its performance. When data input is substantial, as in a patent office, for example, it is a good idea to divorce this operation from the CPU and delegate it to a subsidiary processor. This is the basis for key-to-store input systems. The principle is that data is sent temporarily to some form of auxiliary store (disk, diskette or tape) where it is checked by the subsidiary processor.

The most sophisticated form of this input method is the key-to-disk system. An input station consists of a keyboard and VDU screen, and up to 32 of these can be linked to a minicomputer. This, in turn, is linked to a disk unit.

Data is keyed into a buffer area in the minicomputer while checks are made, and then the data is loaded into the disk system. Each key station can be assigned its own disk space, so the minicomputer is able to support the operations of all of the terminals simultaneously.

The mainframe then gets access to the data on the intelligent terminal's disk either by physically transporting the disk pack to one of the drives used by the mainframe, or the disk can be part of the mainframe disk system. In the latter case, if the accesses of the peripheral processor are not to interfere with the mainframe, the disk system must be **dual ported** – this allows two unconnected computers to write/read the same pack.

Input stations can be distributed around office buildings, in different departments, say, and quite a high throughput of data can be achieved. These systems cost from £15,000 – £35,000, depending upon the number of terminals installed.

Old and new ways of data capture for office systems

Other methods of entering data into office systems includes punched cards and paper tape. However, these methods are more or

less obsolete now and are only used in certain scientific and engineering applications, usually only to program smart machines that the organisation cannot afford to replace.

More modern ways of data entry include **bar coding** and **optical character recognition**; bar codes are at present used mostly for stock control in supermarkets and warehouses. Optical character recognition is still largely experimental, though some prototype systems do exist.

Optical character or optical mark recognition involves making what is written or marked identifiable by a computer. Usually this means that the writing has to be done with a special ink, one based on magnetic iron oxide, for example. This sort of reader is known as a **magnetic ink character recognition system** (MICR). Although MICR is now out of fashion, sophisticated readers are available for standard typing. The system scans characters with light but must be told which typeface is being used.

In some cases, data can be collected automatically, without the need for keying in. This kind of data comes largely from process control applications. Here, data about the quantity and quality of output is sent to some controlling host machine – in a manufacturing plant, for example. Some kind of feedback system may be in operation where the host says 'work faster' to the production line robots, because of data that says not enough products are coming off the line.

Touch-screen technology is now becoming available on personal computers and office workstations (for example, Hewlett-Packard). Two techniques are used:

- 1) **Capacitive** – here, the screen is divided into squares of a large capacitor and changes in the capacitance can be detected when the user presses the screen.
- 2) **Beam-cutting** – this is a more rugged system than the latter, utilising horizontal and vertical infrared beams which the user cuts as the screen is touched and the position is decoded.

Another interesting method of data input is by **voice**. Some tentative steps have been taken towards addressing speech input, but there are considerable problems in programming and training.

Data storage

One of the potential problems with using a small computer to keep business records is that many of the microcomputers available today simply don't store enough information in one place. What kinds of storage device are available for office systems?

Magnetic recording

Magnetic disks don't all hold the same amount of information: a small floppy disk, for instance, usually holds between 150-160 kbyte of information – the equivalent of about 100-105 pages. (However, densities are improving, there are 5¼" floppies available capable of holding 1 Mbyte.)

The records and documents that businesses tend to use become extensive fairly quickly, especially when the computer efficiently handles sorting and filing. It is very easy to go beyond the capacity of small storage media – although disks can be changed to add new information, most programs require access to all the information at once to be efficient. You can, however, add more disk drives by **clustering** them together so that the computer can have more than one disk drive on-line at the same time.

Probably the biggest breakthrough in information storage for small business systems came with the introduction of hard magnetic disks. These, known as Winchester disks after the original development project that created them, contain considerably more storage space and operate much faster than normal floppy disk drives. Typical storage on a Winchester disk is 5-10 Mbyte, well over 4,000 pages of information. All of this data is on-line and the computer can locate any page and load it into main memory in fractions of a second.

Data can be organised into a database, where every stored record can be accessed through certain keys on the keyboard. In a library, for example, it's like having books listed under author, title and subject. Any combination of the facts about the book would locate it in the library. Databases are useful structures for inventory management, customer and client lists, financial records and personnel management.

Digital optical recording

Over the next few years, the capabilities of office information storage systems will be extended by the use of non-magnetic technology. **Optical storage** is the application of laser technology to the problems of information storage and will bring new life to many types of office system computer. They will also add information storage capabilities to classes of system for which magnetic media are not practical.

To overcome the limitations of magnetic disk technology and to push down the cost of randomly (or directly) accessible data, the computer industry is

write part of the magnetic read/write head is replaced by a laser – a photo sensor, required to detect the light, replaces the read head. Data is recorded by burning pits representing binary coded information onto the disk. A low power laser light is used to detect the presence or absence of pits. In effect, this reads the data previously stored. Since the pits cannot be refilled, the recording process is permanent and the drives must be viewed as 'write once/read only' devices. Random access is achieved by disk rotation and radial arm movement, just as in magnetic disk drives.

Using laser light achieves much grea-

Right: PRESTEL service in use in a travel agent's office.



The Research House/British Telecom

turning to the laser as a reading and writing mechanism for disk systems.

Optical recording systems are still in the experimental stages and magnetic media manufacturers have responded to the challenge with the introduction of **vertical** recording media which dramatically increase densities by about ten-fold. (The term vertical because the magnetic domains are aligned perpendicularly to the plane of the disk, rather than in the plane of the disk.)

Optical disk drives perform the same functions as their magnetic equivalents, but in place of the magnetic disk is a disk of material that is sensitive to laser light; the

ter packing densities than magnetic media. Densities of 600 kbit/mm have been demonstrated. This kind of density allows one gigabyte of data (one thousand million bytes) to be stored on one side of a 12" disk.

Another advantage of using laser light is that it can be focused over greater distances than the typical head-to-disk gap of magnetic disk drives. Consequently, as the laser head is further from the disk surface, the drive is not as prone to the dust particles and surface irregularities that cause head crashes on magnetic disk drives. In addition, this means that the disks can be exchanged like hard disk cartridges, but without the contamination



Left: Apple Macintosh uses a 32-bit MC68000 microprocessor. It has 64K bytes of ROM, 128K bytes of RAM and has a built-in polyphonic sound generator capable of producing high quality human speech or music. (Photo: Apple).

problems of cartridges.

Optical technology may be able to reduce the cost of data storage. The use of large head to disk gaps, combined with self correcting alignment systems, simplifies drive engineering. Media costs will also become lower than magnetic media in the future, say the manufacturers.

Further developments for optical disk drives involve improvements in capacity (laboratory systems are claiming 5 Gbyte capacities) and the development of re-writeable systems.

These developments will only benefit office computer systems, though, if they can be accomplished without jeopardising the advantages of optical systems, specifically the low error rates required for true digital applications. These problems are likely to keep these systems in the laboratory for some time, so updateable optical media are unlikely to become practical for office applications until the 1990s.

More promising is the development of small optical disk drives of 5 inch diameter or smaller media. These are more

likely to form the next generation of word processor storage devices, if only to reduce the large floppy disk archives building up within every word processing pool.

The applications of such optical disk systems fall into three areas: back-up systems, high capacity data stores and image storing systems.

The optical medium compares fairly well with the cost of $1,600 \text{ bit in}^{-1}$ tape, though tape is reusable. Optical disk drives are more expensive, but they do offer significant advantages in terms of convenience of media handling and of random access for selective system restore.

High capacity disk storage is becoming more important as the growth of office systems generates more business information in a digital form. One optical disk represents the contents of 1,000 floppy disks or 500,000 A4 pages of word processor output.

Optical recording can be seen as delivering the archival storage for the file servers of office automation systems of the future.

Communications in the office

Electronic communications are most commonly known as telecommunications: the use of telephone, telegraph, radio, television etc., to transmit information. The information transmitted in a telecommunications system may be speech, music, pictures or data depending on the application. Data communications around the office and between offices may use a combination of telecommunications methods.

Information moves very quickly through any computer system. At the centre of this movement is the main memory of the computer, which holds the program telling the computer what to do. The program pulls the information into main memory before it is either manipulated or sent somewhere else. When creating or editing a report at the office microcomputer, say, the editing program and the information that is being typed are located in main memory. The computer first sends the information from the memory to the screen to be displayed. When a sentence is changed the program makes the change in main memory first, then it sends the information to the screen for the user to see. When the report needs to be stored, the computer copies the information that is in main memory onto a disk.

The report doesn't have to stay in disk storage; if a printout is required the program copies the report from the disk into main memory a section at a time, and sends it from there to the printer.

In the same way that information can be sent around an office system, information can also be sent to another computer: if the second computer is in the same building, the information is sent directly over a special communications cable; if the second computer is not in the building, the two can be linked across telephone lines using a **modem** at each end. Modem stands for modulate/demodulate. The modem converts digital computer signals into analogue noise for sending down a telephone line. At the other end, another modem demodulates the noise, converting it into digital pulses which the second

computer can understand.

Modems plug directly into a telephone line at ordinary British Telecom sockets. A related device, an acoustic coupler, links computer devices to the telephone directly. An acoustic coupler looks like a double egg cup and the telephone handset rests in the two receiving cups. This is a relatively inexpensive and convenient way of connecting small microcomputers to larger office systems and forms the basis for all communications systems. However, because the telephone is connected acoustically, instead of being wired to the computer, loud background noises can sometimes penetrate the acoustic seal, thus causing transmission errors.

Acoustic couplers transmit data at either 300 or 1,200 bit s⁻¹, which is fast enough for most simple office applications. In comparison, modems can transmit data at up to 9,600 bit s⁻¹ over standard voice grade telephone lines, but slower speeds are more common.

Unlike acoustic couplers, modems are wired directly to digital computer devices and the transmission line. This limits portability, but prevents extraneous room noise from interfering with transmission and reception.

Some modems can only dial out, like certain payphones, but others, known as auto answer modems, can also automatically answer and connect calls to the local parent device. Some modems of both types have telephones attached for placing and receiving normal calls.

Users rarely tax the full capacity of the transmission medium, and this is obviously inefficient. Some methods are available which effectively divide the transmission line so that more than one user may use the same link. This is known as **multiplexing**.

Frequency division multiplexing is one type of multiplexing method which functions rather like a radio communications system. Information comes over the air waves (or transmission line) and when it arrives, the various users tune into the right frequency and get the signals that are meant for them.

Time division multiplexing, in contrast, allows each computer device to send information down the line a little at a time

in strict rotation. This **polling** occurs so quickly that it appears that all users have their own circuit. Both of these multiplexing methods are explained in greater detail with regard to telephone systems in *Communications 3*.

Although a terminal or computer is inactive at certain times, time division multiplexing continues to give it time to transmit. Frequency division multiplexing similarly apportions sections of the frequency bandwidth to terminals, regardless of whether they need it or not. We can see, therefore, that for a proportion of the total time, nothing is transmitted in some (or even all!) of the respective time or frequency slots. Often, a whole transmission link may be full of slots, capable of carrying information, with *nothing in them* – an inefficient state of affairs. This prompted the design of the **statistical multiplexer** to enable the transmission of data more efficiently.

A statistical time division multiplexer, to give it its full name, will bypass a momentarily inactive terminal in a polling sequence and give the transmission time to the next terminal waiting with data. To the users it seems as if the multiplexer is polling only the active terminals, which consequently get more time for transmitting; a direct statistical relationship – more data, more transmission time.

Communications networks

Communications networks transmit digital data and analogue voice or both signals at the same time. Although telephone companies are the largest carriers of this data, many other companies switch data into the public telephone network, or use leased circuits, because their networks are so large. Collectively, these networks are called **wide area networks**.

Local area networks, on the other hand, are to link equipment that is contained wholly on one site. They do not require the public authorities or anyone else to be involved.

To make the most of local area networks, provision should be made for linking them into wide area networks. This is usually accomplished by way of special computers called **gateways**, which perform protocol and speed conversion and

which act as an interface between the networks.

Protocols are technical customs or guidelines that govern the exchange of signal transmission between equipment. Each protocol specifies the exact order in which signals are to be transferred, what signal will indicate that the opposite device has completed its transfer, and so on.

Hardware and software are both designed to handle specific protocols and the protocols themselves are often named after the device for which they were designed. Both wide area and local area networks use specific protocols, e.g. British Telecom's X.25 for the former and RS-232C for the latter.

Only devices using the same protocol can communicate directly with one another. Devices using different protocols must transmit and receive data through an intermediate device such as a protocol converter (in hardware) or an interpretation program (in software). Protocol can be thought of as the red tape of data communications.

Topology of networks

To meet many requirements, computer networks have developed in a number of ways. A centrally based network with links to devices feeding it with information and allowing users access to the results is based upon a **star** type system (*figure 1a*).

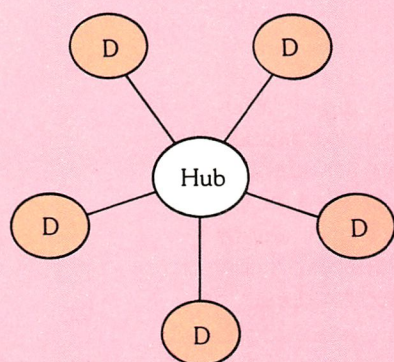
Another common design is the **mesh network** (*figure 1b*), in which a number of nodes are connected together. Information from one node sometimes has to pass via a third party to reach its destination because no direct link exists. There is also the fully interconnected network, in which every node is connected to every other one.

Less common at present, but increasing in importance, are the **bus** and **ring** or **loop** networks (*figures 1c and d*).

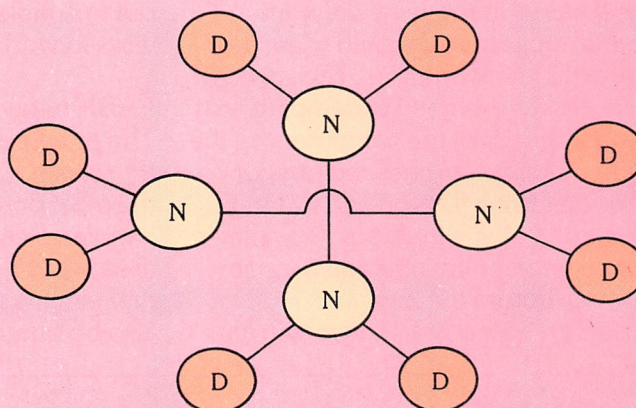
To fulfill the different requirements of office users, manufacturers and suppliers of network equipment have produced various facilities such as **multiple routing**.

Two schemes have evolved here. The first, **alternative routing**, describes the situation where information normally follows one route between any two nodes, but it can also be made to take another if the network (or any link) fails or is over-

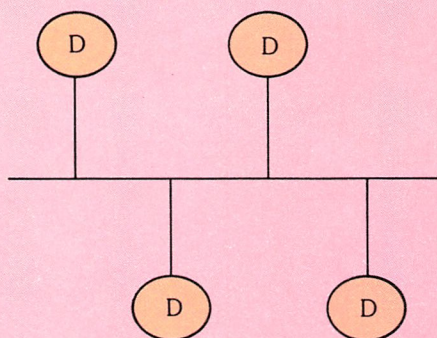
1



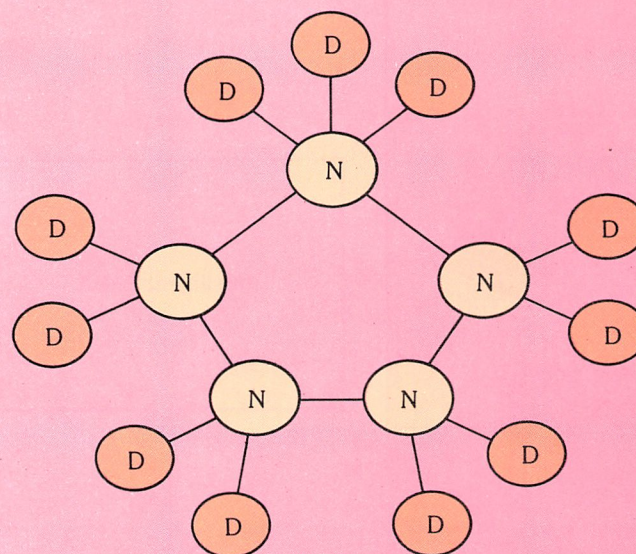
a) Star (or hub)



b) Mesh



c) Bus



d) Ring

where N = Node (computer or controller)
D = Automation device

1. Common designs for networks.

loaded. The second system is **adaptive routing**, where a network management system keeps track of the state of traffic on every line and decides at the time of transmission of a message which route it should take.

These systems enable a network to continue functioning even when links or nodes fail. Afterall, the only concern of automated office users is that data is passed from A to B. They don't need to know how it gets there.

Another facility often found on data networks is **file transfer**, but this can be difficult if the source and target systems are not matched, in terms of data transfer rate, for example. In addition to moving files, users usually want to be able to manipulate

them (update, delete, copy, archive etc.) from a remote location. A significant feature of the more traditional computer systems is the rather slow data transfer rate that they support. A rate of a few thousand bits per second can severely hamper data and file transfer between computers, although it may be more than adequate for operator-terminal transactions.

Linking computers by circuits that operate at speeds significantly lower than the operating speeds of the computers can cause traffic difficulties. This is so unless the systems themselves have made provision for such links in the form of communications hardware and software.

The **Cambridge ring network** was designed for use within the University of

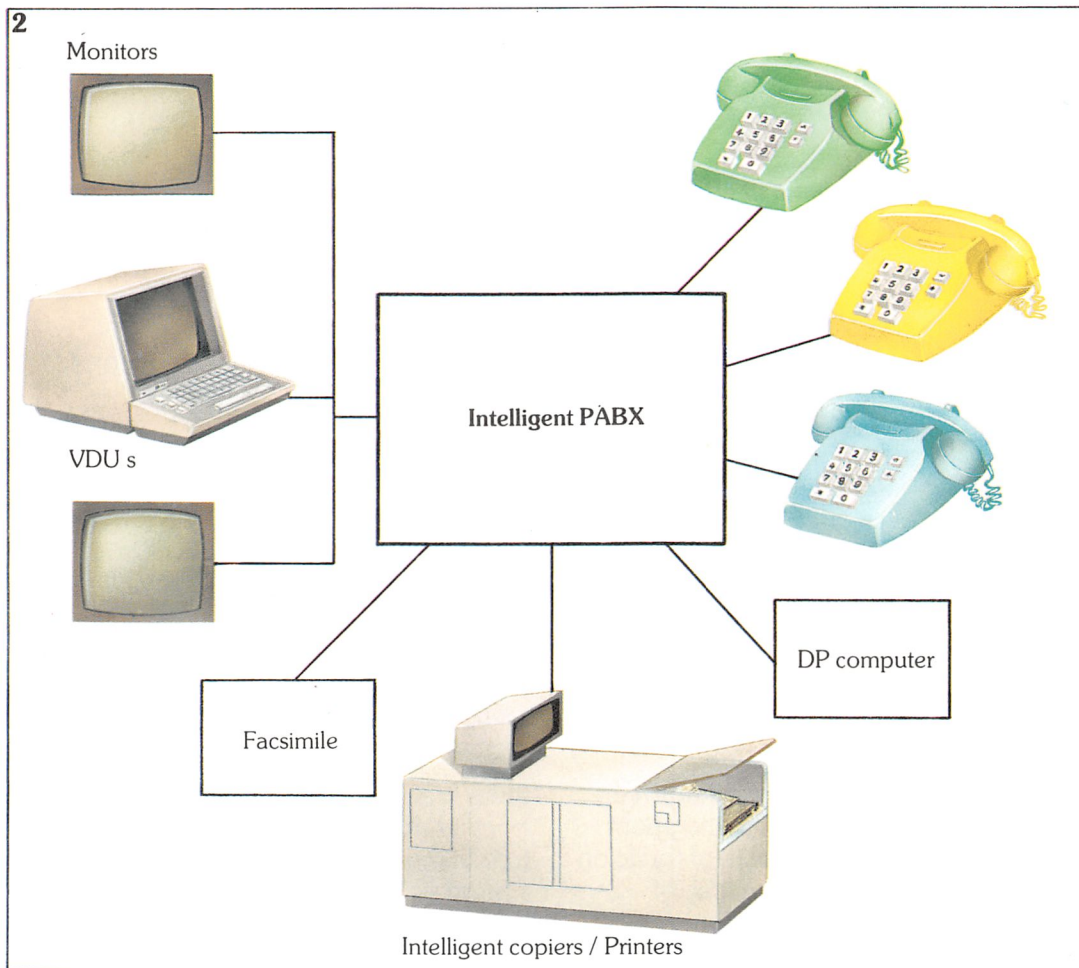
Cambridge computer laboratory. The ring was designed so that several different computers and terminals could share a range of equipment and files. The speed of the network was designed to compare with the internal speeds of the computers on the ring and 10 Mbit/s was chosen as the raw transmission speed. This gives a transfer rate of about 1 Mbit/s from point to point in practice because of the other data traffic also flowing around the ring.

Installing a local area network in an

tocol and code conversions can be carried out to enable otherwise incompatible devices to work together.

PABX networks

The private telephone exchange (PABX, for private automatic branch exchange) is the most obvious candidate for adapting to handle voice and data traffic on a local area network. The same telephone lines that join the telephone to the PABX can also be simultaneously used for data. The



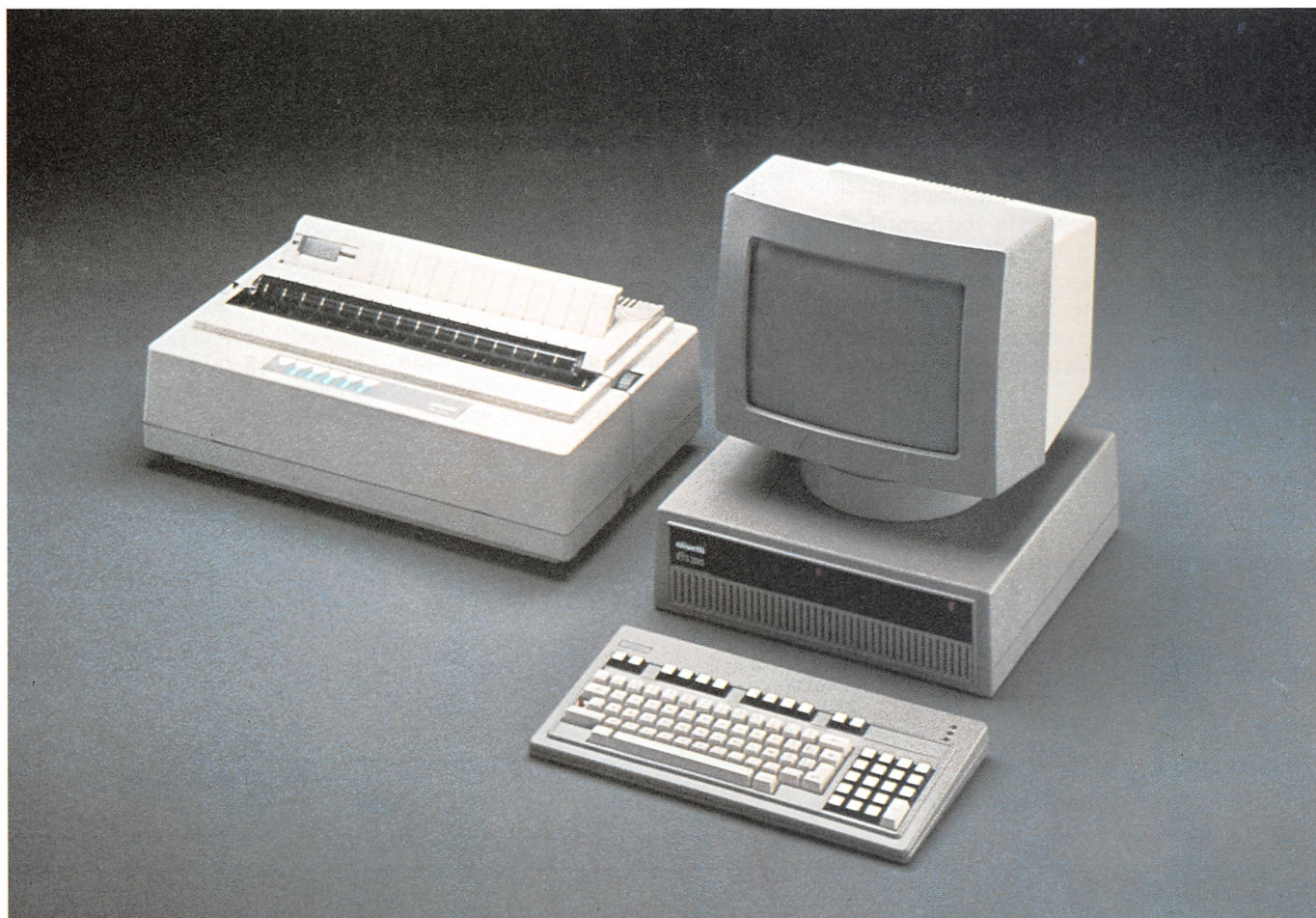
2. A possible office network based on the PABX.

office block, using the corridors for the main ring wire, can simplify many of the wiring problems of using multiplexers and other line concentrating types of equipment.

Once the network is added to an office building, new services and features can be added to enhance the system. Usually a mailbox and shared file facilities can be designed, and providing the correct network interface units are available, pro-

PABX differentiates between the two channels and routes them accordingly (figure 2).

With such systems, no new wiring is needed since telephone lines of adequate quality are available in most offices. However, the idea does have some disadvantages. For example, telephone lines are unable to handle high transmission speeds, which makes them unsuitable for bulk data transfer or for use with devices that incorporate microprocessors. In addi-



Above: Olivetti ETS 2010. (Photo: Olivetti).

tion, computers need to be talking with several devices at the same time and this can be difficult to organise in what is really just a circuit switching device.

Besides the fact that new wiring is not needed, the PABX approach can have other advantages. For example, the PABX needed to handle mixed data and voice will also be able to handle the more advanced telephone routing functions: multiple users involved in calls (the precursor of teleconferencing, which we'll be looking at in the next chapter); automatic redirection; and hold on to busy numbers and call them again when free.

The cost of a new PABX to provide these facilities can be often justified for speech alone. Also, the advantages are more easily demonstrable than for a new local area network.

In the fully integrated electronic office, voice and data communications will both be handled. However, it is not yet clear whether the PABX or the shared

resource local area network will provide the communications medium. One possibility is that the PABX will become just a node on the local area network, thus allowing information to pass between the two of them.

Conclusion

Local area networks, in whatever form they finally take, are a vital element in the linking together of the components we have discussed. In particular, they provide the user at a workstation with access to all other workstations and services that can be provided by a shared resource system.

Local area networks can provide the catalyst for the integration of different office services and it can also be the medium for office communications. These include voice messaging, word processing and facsimile devices that transfer images. It is these, the functions of office automation systems, that are the subject of the next chapter.

ELECTRICAL TECHNOLOGY

Apparent power

The concepts of real (or active) and reactive power are summarised in the power triangle in figure 1. They may be related to the **series circuit model** shown in figure 2a, where an impedance, Z , is represented by the combination of a resistance, R , in series with a reactance, X , which may represent either a capacitor or an inductor. The voltage and current phasor diagram of this may be drawn as in figure 2b, where I is the reference phasor. Now we may write:

$$V_R = V \cos \phi$$

$$= RI$$

and:

$$V_X = V \sin \phi$$

$$= XI$$

Here, we see that the voltage is resolved into two components: V_R , in phase with the current; and V_X , in phase quadrature. By multiplying each of these components by the current, we obtain for the real power:

$$P = V_R I$$

$$= RI^2$$

$$= V \cos \phi I$$

Similarly, for the reactive power we get:

$$Q = V_X I$$

$$= XI^2$$

$$= V \sin \phi I$$

Parallel model

We have seen that it is also possible to construct an alternative circuit model for any device comprising a conductance of value G , in parallel with a susceptance of value B , as shown in figure 3a. The phasor diagram for this circuit is shown in figure 3b where we see that, in this model, the current is resolved into the two components: I_G , in phase with the voltage; and I_B , in phase quadrature. Thus:

$$I_G = I \cos \phi$$

$$= VG$$

and:

$$I_B = I \sin \phi$$

$$= VB$$

Multiplying each of these by the supply voltage we obtain:

$$P = VI_G = VI \cos \phi$$

and:

$$Q = VI_B = VI \sin \phi$$

which is *exactly the same as before*. This, in fact, shows us that no matter how a circuit is modelled, the *values of real and reactive power are unchanged*.

Apparent power

The **apparent power**, S , in a circuit is obtained

(figure 1) by summation of the two phasors representing real and reactive power, where:

$$S = \sqrt{P^2 + Q^2}$$

We may therefore write:

$$P = S \cos \phi$$

and:

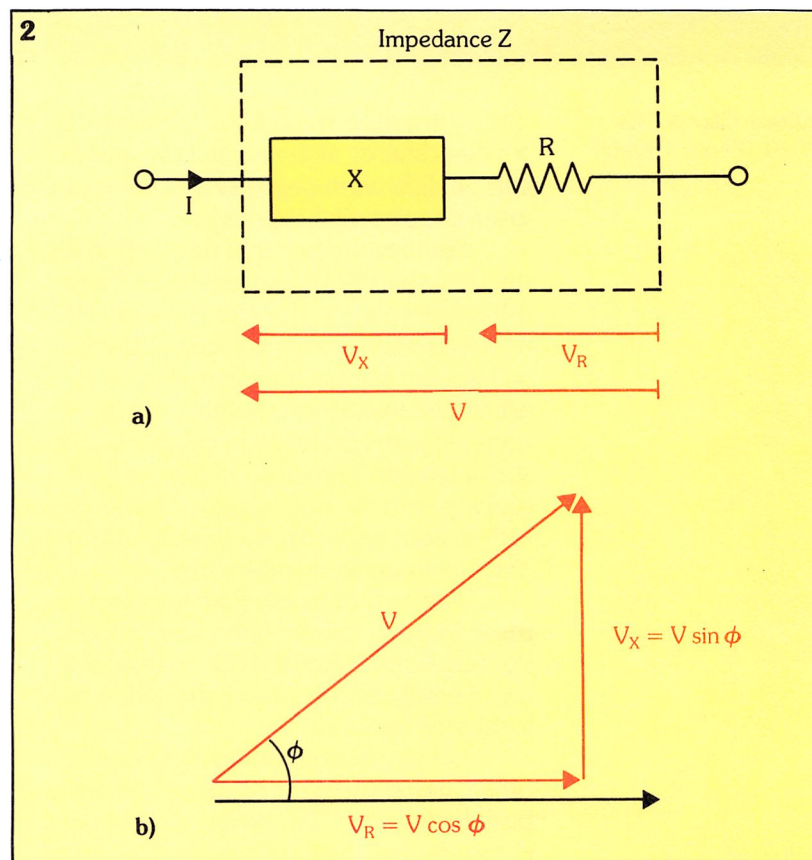
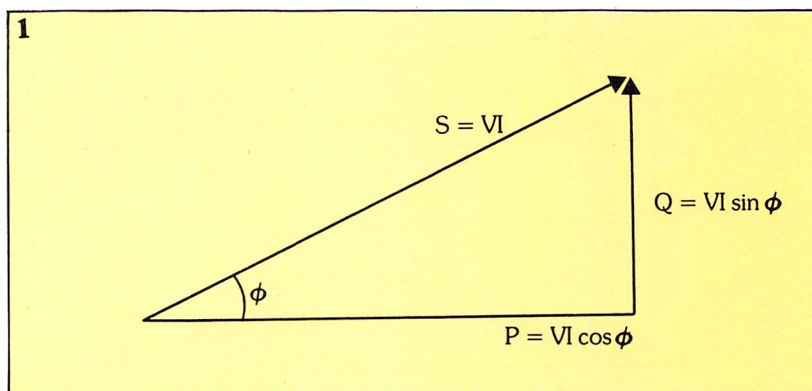
$$Q = S \sin \phi$$

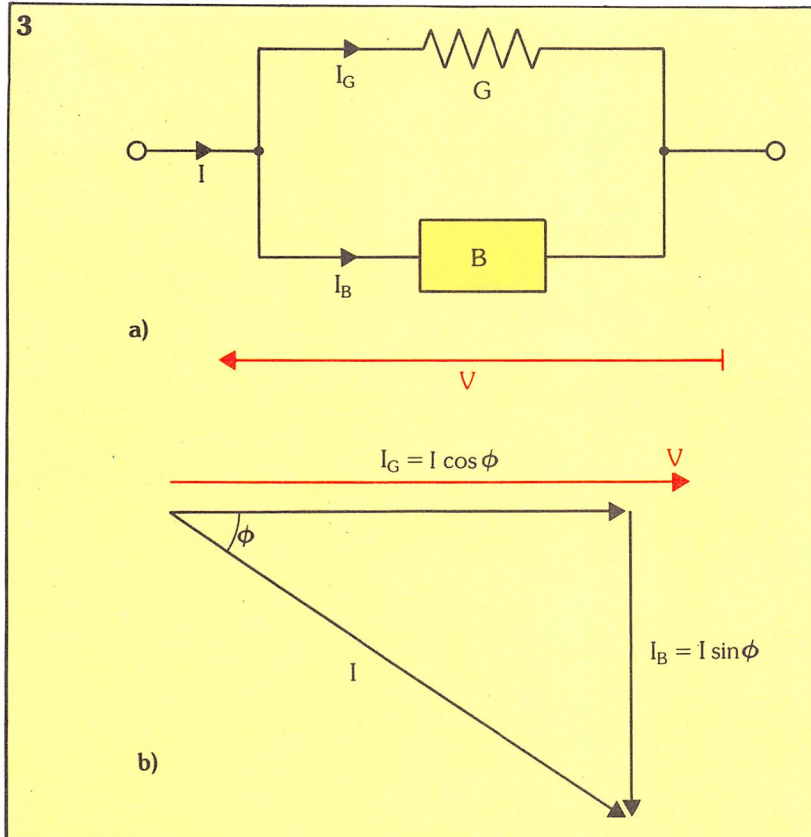
Apparent power in parallel circuits

This ability to resolve the apparent power into the two components of real and reactive power

1. Power triangle.

2. (a) Series circuit model;
(b) voltage and current phasor diagram.





3. (a) Parallel circuit model; (b) phasor diagram.

is useful when a number of circuits are connected in parallel to a single supply and the total load taken from the supply needs to be determined.

The significance of this can be seen when recalling the number of appliances in a single house: lights, heaters, cooker, refrigerator, television, water heater, etc., all of which are connected in parallel to a single supply. Furthermore, homes, factories, offices, etc., each taking a separate load, are all connected in parallel to the National Grid.

Consider a set of loads with apparent powers S_1, S_2, S_3 , and power factors $\cos \phi_1, \cos \phi_2, \cos \phi_3$, and we wish to determine the total load from the supply. The power phasor diagram is shown in figure 4a, where each apparent power is drawn at its appropriate angle and placed head to tail, so that they can be added together. The resultant, S , of these three phasors can be drawn from the starting point to the end of the chain of phasors, as shown.

The easiest method of calculating the value of this resultant total apparent power is to resolve each load into its real and reactive power, as shown in figure 4b. Thus for load S_1 we have the two components:

$$P_1 = S_1 \cos \phi_1$$

and:

$$Q_1 = S_1 \sin \phi_1$$

and similarly for all the other loads in the system. The real and reactive components of the total load on the supply can be found by adding the appropriate components of the individual loads, as shown in figure 4b. This gives:

$$P = P_1 + P_2 + P_3 + \dots$$

and:

$$Q = Q_1 + Q_2 + Q_3 + \dots$$

Note that the sign of the reactive power must be taken into account: considering inductively reactive power as positive and capacitive power as negative. In the example shown, real power P_3 has a negative power factor and so reactive power Q_3 is a negative quantity.

From this, the total apparent power S can be derived from:

$$S = \sqrt{P^2 + Q^2}$$

and its **power factor** from:

$$\cos \phi = \frac{P}{S}$$

The power factor will be positive if the inductively reactive power is greater than the capacitive element. We can show the calculations of apparent power and power factor with a couple of examples.

Example 1

A capacitive and inductive circuit is shown in figure 5. The real power, P , in the complete circuit is the sum of the powers P_1 and P_2 in each branch:

$$\begin{aligned} P &= P_1 + P_2 \\ &= I_1^2 R_1 + I_2^2 R_2 \\ &= (2^2 \times 10) + (1.41^2 \times 30) \\ &= 40 + 60 \\ &= 100 \text{ W} \end{aligned}$$

The reactive power is, similarly:

$$\begin{aligned} Q &= Q_1 + Q_2 \\ &= I_1^2 X_L + I_2^2 X_C \\ &= (2^2 \times 20) + (1.41^2 \times 10) \\ &= 80 + 20 \\ &= 100 \text{ W} \end{aligned}$$

The total apparent power is given by:

$$\begin{aligned} S &= \sqrt{P^2 + Q^2} \\ &= \sqrt{100^2 + 100^2} \\ &= 141 \text{ W} \end{aligned}$$

and the power factor is:

$$\begin{aligned} \cos \phi &= \frac{P}{S} = \frac{100}{141} \\ &= .707 \end{aligned}$$

Example 2

Consider a small domestic installation comprising:

- a) an electric cooker taking 2 kW at unity power factor;
- b) a total lighting load of 500 W at unity power

factor;

c) an electric fan working from 250 V and taking a current of 1.5 A at a power factor of 0.8;

d) a washing machine taking an apparent power of 2 kVA, at a power factor of 0.7.

Listing the real and reactive powers for each item we have: cooker:

$$P_1 = 2000 \text{ W}; Q_1 = 0 \text{ W}$$

lights:

$$P_2 = 500 \text{ W}; Q_2 = 0 \text{ W}$$

fan:

$$\begin{aligned} P_3 &= VI \cos \phi_3 \\ &= 250 \times 1.5 \times 0.8 \\ &= 300 \text{ W} \end{aligned}$$

$$\begin{aligned} S_3 &= VI \\ &= 250 \times 1.5 \\ &= 375 \text{ VA} \end{aligned}$$

$$\begin{aligned} Q_3 &= \sqrt{S_3^2 - P_3^2} \\ &= \sqrt{375^2 - 300^2} \\ &= 224 \text{ VAR} \end{aligned}$$

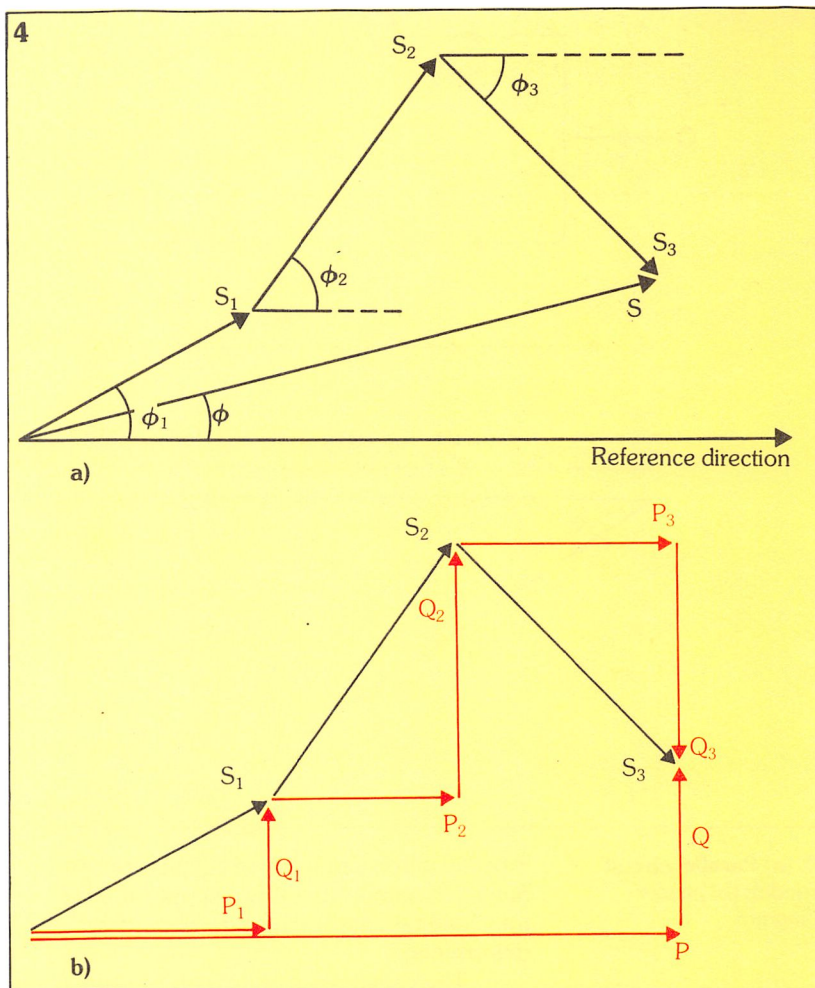
washing machine:

$$S_4 = 2000 \text{ VA}$$

$$\begin{aligned} P_4 &= S_4 \cos \phi_4 \\ &= 2000 \times 0.7 \\ &= 1400 \text{ W} \end{aligned}$$

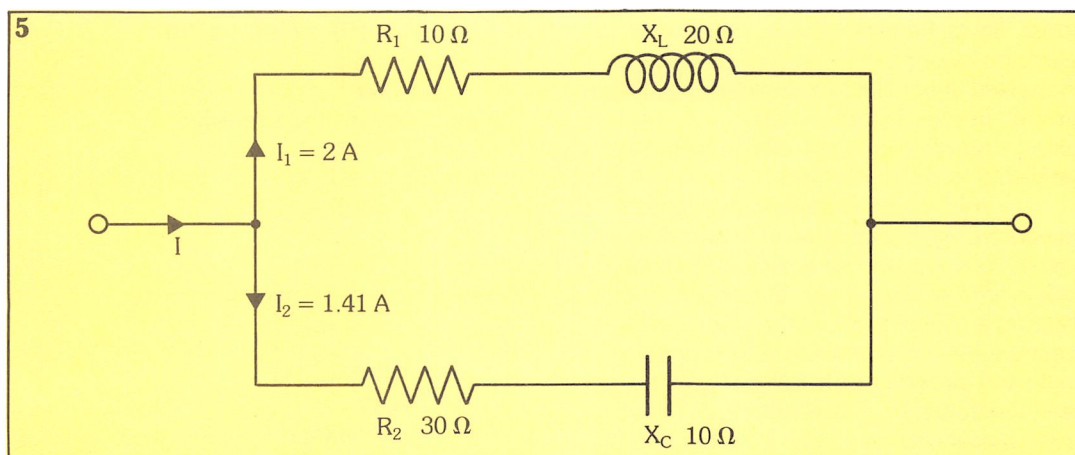
$$\begin{aligned} Q_4 &= \sqrt{S_4^2 - P_4^2} \\ &= \sqrt{2000^2 - 1400^2} \\ &= 1430 \text{ VAR} \end{aligned}$$

We may now compute the total real and



4. (a) Power phasor diagram for the set of loads S_1 , S_2 and S_3 ; **(b)** resolving each load into its real and reactive power to determine the apparent power.

5. A capacitive and inductive circuit.



reactive power, where real power is given by:

$$\begin{aligned} P &= P_1 + P_2 + P_3 + P_4 \\ &= 2000 + 500 + 300 + 1400 \\ &= 4200 \text{ W} \end{aligned}$$

and reactive power is given by:

$$\begin{aligned} Q &= Q_1 + Q_2 + Q_3 + Q_4 \\ &= 0 + 0 + 224 + 1430 \\ &= 1654 \text{ VAR} \end{aligned}$$

The apparent power is:

$$\begin{aligned} S &= \sqrt{P^2 + Q^2} \\ &= \sqrt{4200^2 + 1654^2} \\ &= 4510 \text{ VA} \end{aligned}$$

and the power factor of the total load is:

$$\cos \phi = \frac{4200}{4510} = 0.931$$

□

Analogue circuits-2

Oscillators

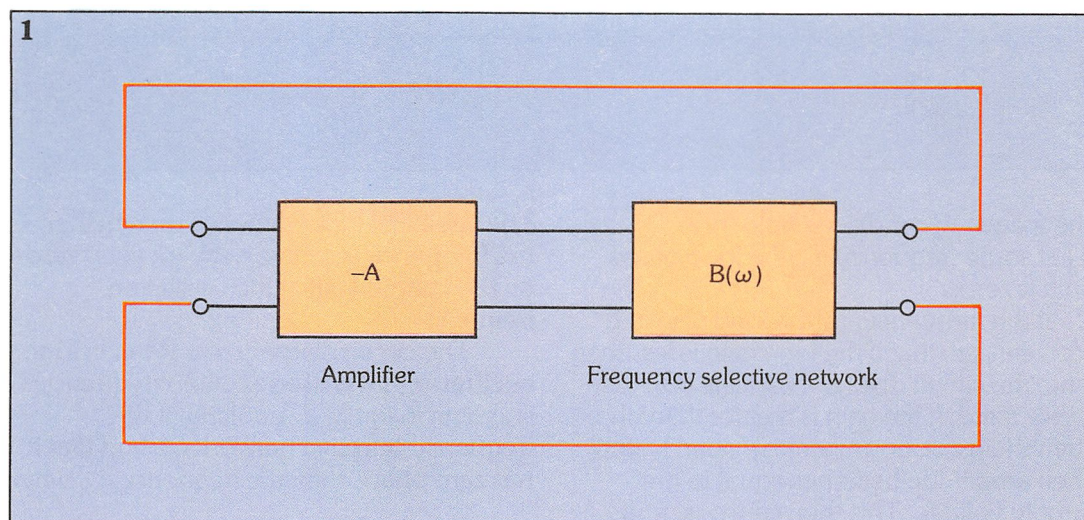
Oscillators form a very large group of analogue circuits. We looked at some oscillator circuits briefly in earlier chapters, but let's now concentrate on oscillators that generate sine waves. A sinusoidal oscillator is defined as a circuit that will generate a sine wave of a chosen frequency without needing any signal input.

The operating principle behind an oscillator can be seen by considering figure 1, where an amplifier followed by a

$+180^\circ = 360^\circ$ (which is equal to zero) while the magnitude of the whole circuit's gain is AB_o .

If we now adjust the amplifier's gain so that $AB_o = 1$, the output of this system at frequency f_o is exactly the same as the input. At any other frequency, the output is not in the same phase. So if we were to connect a variable frequency signal generator to the input of the amplifier, the output signal from the network would have zero phase shift when the generator was set to f_o and at no other frequency.

1. Operating principle behind an oscillator.



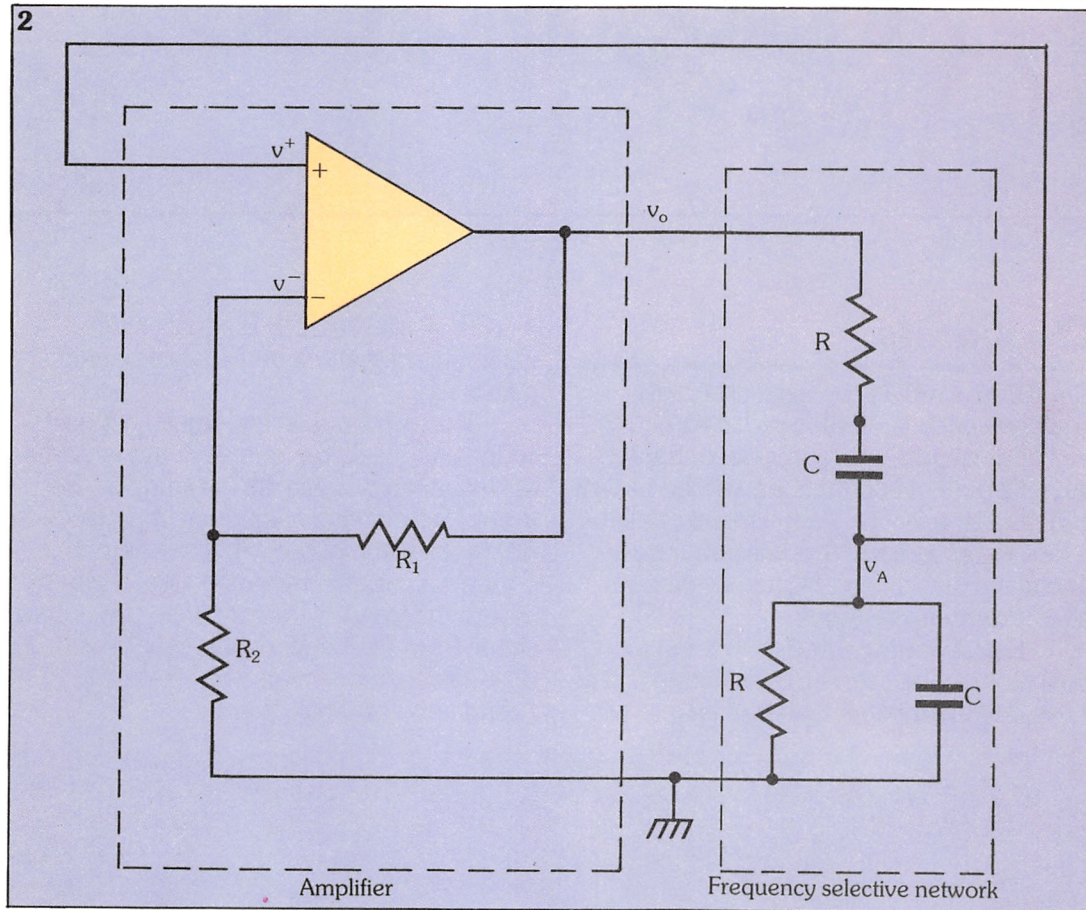
frequency selective network (or filter) is shown. The amplifier has a voltage gain of A and a phase shift of 180° (indicated by the negative sign in the diagram). The frequency selective network has a transfer ratio represented by $B(\omega)$. Written in this way, it shows that both the magnitude and phase of the ratio are functions of frequency.

If we assume that at a particular frequency, f_o , this network's phase shift is 180° and the voltage gain (more likely to be a loss) is B_o , then the overall phase shift from the amplifier's input to the frequency selective network's output is given by 180°

If the signal generator is removed from the circuit, and the output fed back to the input (as shown in red), then the network will oscillate sinusoidally, as a component of frequency f_o is supplied by the amplifier's inherent noise which causes the oscillation to be set up initially (see *Solid State Electronics 24*). The same thing would of course happen, if the amplifier had a gain of $+A$ (zero phase shift) and if the frequency selective network also had zero phase shift.

This is summarised in the **Bar-khausen conditions for oscillation**, which state that a network will oscillate if both of

2



2. Circuit diagram of a Wien bridge oscillator.

the following conditions hold true:

1) at some frequency, f_o , the loop phase shift is zero;
 2) at the frequency, f_o , the loop gain is 1.
 You can see that if the loop gain is less than one, the circuit does not oscillate. On the other hand, if the gain is greater than one, then the oscillations increase steadily until their amplitude becomes equal to the supply voltage. This then causes serious distortion that flattens the peaks of the sine wave, thus creating a square wave. However, to prevent this happening, the gain of the amplifier is arranged so that it decreases when it reaches saturation, so the oscillator has a controlled amplitude and is kept almost perfectly sinusoidal.

Wien bridge oscillator

All oscillators involve an amplifier, which is usually an op-amp. In *Solid State Electronics* 24 and 25, we saw that by utilising the non-inverting input op-amps are able to provide gain with zero phase shift, or alternatively by using the inverting input, gain with 180° phase shift can be achieved.

The choice of input depends on whether the frequency selective network gives zero or 180° phase shift at the oscillation frequency.

The circuit diagram of a **Wien bridge oscillator** that is used at audio frequencies is shown in figure 2. Looking at the frequency selective network we find that it has zero phase shift at a frequency, f_o , where:

$$f_o = \frac{1}{2\pi CR}$$

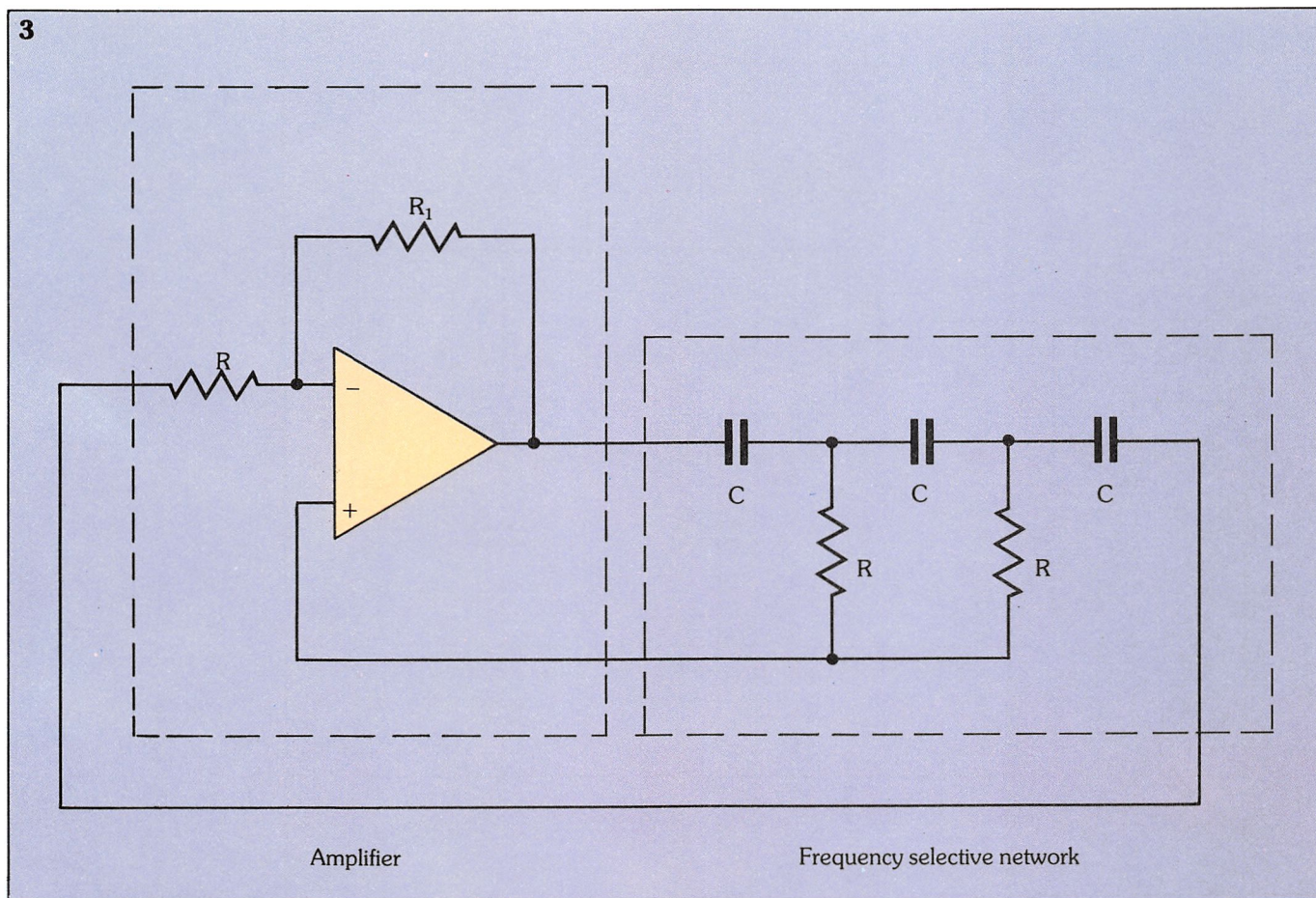
At this frequency, the voltage gain of the network is:

$$\frac{v_A}{v_o} = \frac{1}{3}$$

So you can see that for oscillation to occur the amplifier must be designed to have a gain of 3, so that the loop gain is exactly unity.

We can assume that the gain of the op-amp is very large (feedback being supplied by the resistors R_1 and R_2). This being the case, we may say that the voltage between the two input terminals is approximately zero. Thus:

3



3. Circuit diagram for a phase shift oscillator.

$v^+ = v^-$
Furthermore, the voltage V^+ is derived from the output voltage V_o by a potential divider, so:

$$\frac{v_o}{v^-} = \frac{v_o}{v^+} = \frac{R_1 + R_2}{R_1}$$

and as we want the gain to equal 3:

$$R_2 = 2R_1$$

Finally, to ensure that the amplifier's gain falls as the amplitude increases, the resistor R_2 is replaced by a **thermistor** (a device whose resistance has a negative temperature coefficient). As the rms value of the output voltage increases, so does the heat dissipated in R_2 . As its temperature rises, its resistance falls, thus reducing the gain of the amplifier circuit.

Phase shift oscillator

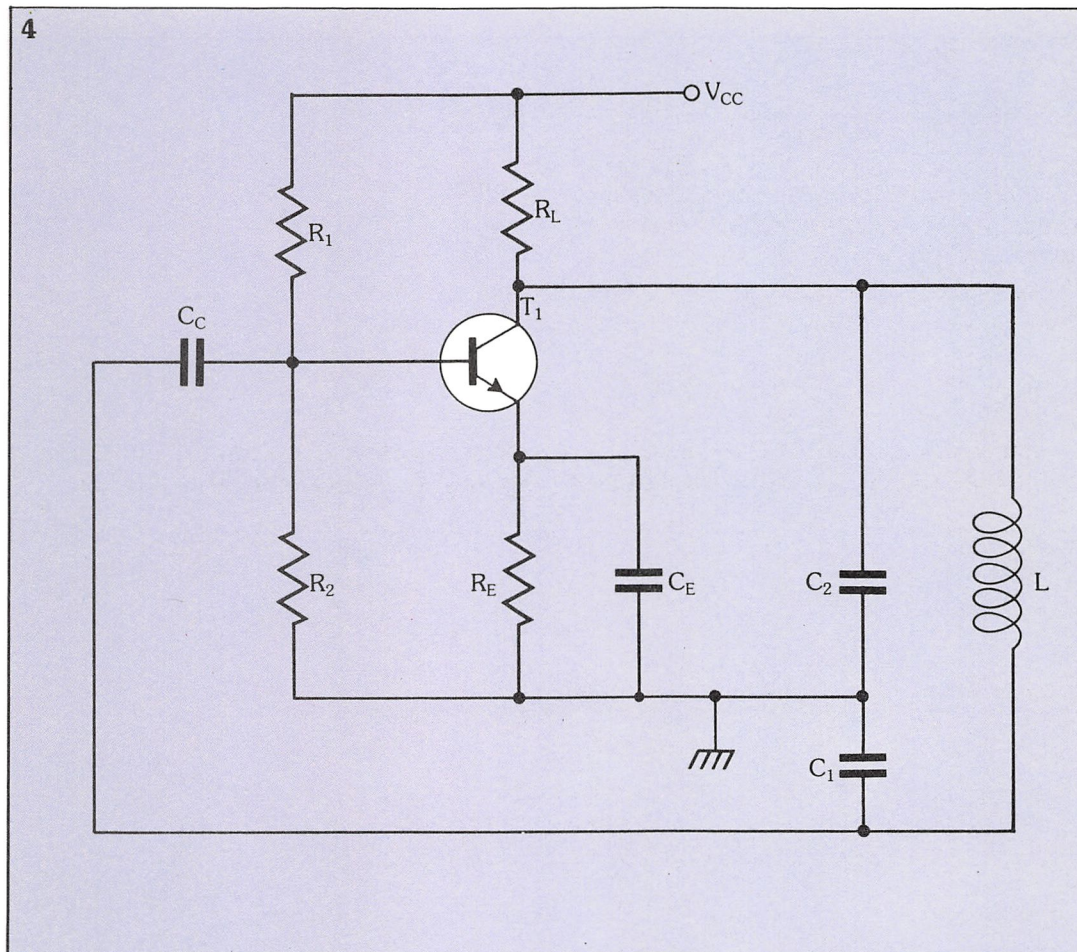
Phase shift oscillators are very common circuits and the diagram for one form of this is shown in figure 3. The frequency controlling network has a phase shift of 180° at a frequency, f_o :

$$f_o = \frac{1}{2\pi RC\sqrt{6}} = \frac{0.065}{RC}$$

At this frequency, the network has a gain of $1/29$; it is therefore necessary to design an amplifier with a gain of 29. This means that the value of resistor R_1 must be at least 28 times the value of resistor R . In actual fact, it is usually made about 5% larger, to ensure that the oscillation will start.

Resistor R_1 may be replaced with a **sensistor**, which is a device with a positive temperature coefficient. As the voltage at the output of the op-amp increases, so the value of the sensistor also increases, thus reducing the gain and stabilising the amplitude.

You'll notice that the feedback connection in this oscillator is taken to the op-amp's inverting input, so as to give 180° phase shift, ensuring that the total phase shift around the complete loop is zero.



4. Colpitts oscillator.

The Wien bridge and phase shift oscillators are useful over the frequency range 10 Hz to 200 kHz. They can also be used to give a variable frequency output by varying the value of capacitor, C . It is only necessary to vary one of the capacitors in either circuit, but more usually all the capacitors are made variable and mounted on a single shaft, so they can all be altered together. This gives us the maximum frequency range, limited only by the design of variable capacitors which have a maximum to minimum ratio of about 10 to 1.

Colpitts oscillator

Figure 4 shows a type of circuit that is suited to high frequency operation – the **Colpitts oscillator**. The frequency selective network in this circuit consists of the parallel resonant circuit comprising the inductor, L , and capacitances C_1 and C_2 in series. This sort of circuit has a zero phase shift at the resonant frequency, f_o , given by:

$$f_o = \frac{1}{2\pi\sqrt{LC}}$$

where:

$$C = \frac{C_1 C_2}{C_1 + C_2}$$

being the series combination of C_1 and C_2 .

The circuit's amplification is, as we have said, supplied by the single transistor T_1 , with load resistor R_L , emitter biasing resistor R_E , bypass capacitor C_E , and base biasing resistors R_1 and R_2 . The capacitor C_C prevents the DC supply voltage from reaching the base and is usually of quite a high value allowing the sinusoidal voltages to pass easily.

At the resonant frequency, the gain of this amplifier is approximately $g_m R_L$, where g_m is the transconductance of the transistor. The voltage gain of the frequency selective network is $C_2/(C_1 + C_2)$, so we need to fix the value of the load resistor R_L so:

$$(g_m R_L) \frac{C_2}{C_1 + C_2} = 1$$

or, in practice, slightly greater than unity.

Op-amps used as integrators

An integrator is a circuit whose output is the time integral (usually abbreviated to integral) of the input wave. So, how do we define the term integral? Figure 5a illustrates a voltage wave, $v_i(t)$. The value of the integral at some time t_1 , $v_o(t_1)$, is obtained by taking the area under the curve up to time t_1 , as shown by the shaded area. If this is done for every instant of time, we shall obtain the complete integral of $v_i(t)$ which is shown by the curve

black) given by:

$$v_i = V_i \cos \omega t$$

The integral of this is a sine curve:

$$v_o = V_o \sin \omega t$$

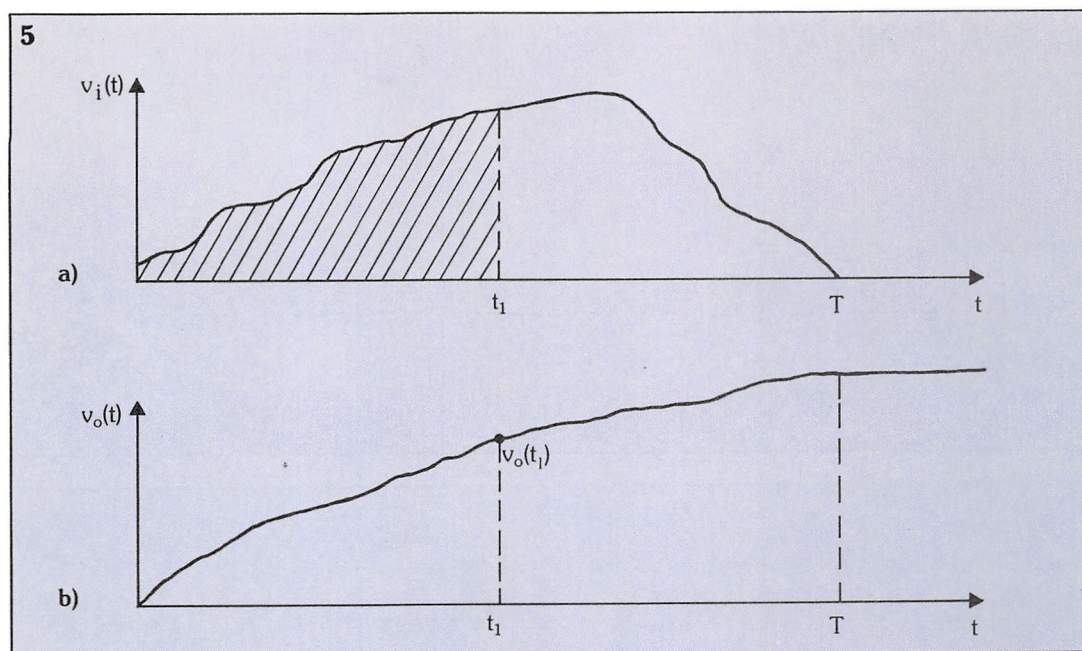
where:

$$V_o = \frac{V_i}{\omega}$$

We can see that the integral of a sinusoidal wave is another sinusoid which lags the first by 90° , and whose amplitude is smaller by a factor $1/\omega$.

An integrator can be constructed very simply by using an op-amp as shown in figure 7. Let's look at the way this circuit

5. Determining the integral of $v_i(t)$.



in figure 5b. Notice that the input voltage is limited to a finite period of time T , and so its integral will steadily increase until time T , after which it will be constant.

We may examine the integrals of some particular voltages. First let's look at a sudden step in voltage as shown by the black curve in figure 6a. When $t = 0$, the area under the black curve is zero so the integral $v_o(t)$ (shown by the red line) starts at zero. As time increases, the value of integral $v_o(t)$ rises linearly with time, resulting in a graph with a constant slope – termed a **ramp**. Of course this ramp cannot continue indefinitely in a practical circuit, and will be cut off by the value of the supply voltage.

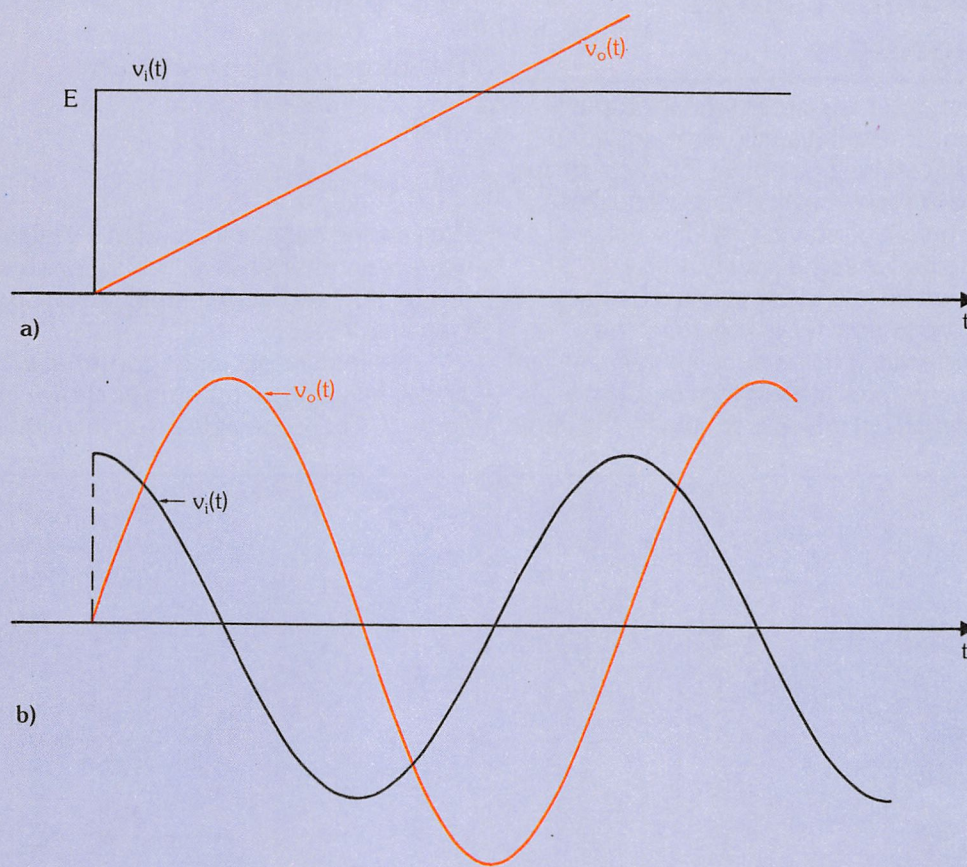
Figure 6b shows a cosine curve (in

behaves when the input voltage takes a sudden step from zero to a value E (as in figure 6a). Since the inverting input in an ideal op-amp is at the same potential as the non-inverting input – in this case at ground potential – the current I , flowing through R , is given by:

$$I = \frac{E}{R}$$

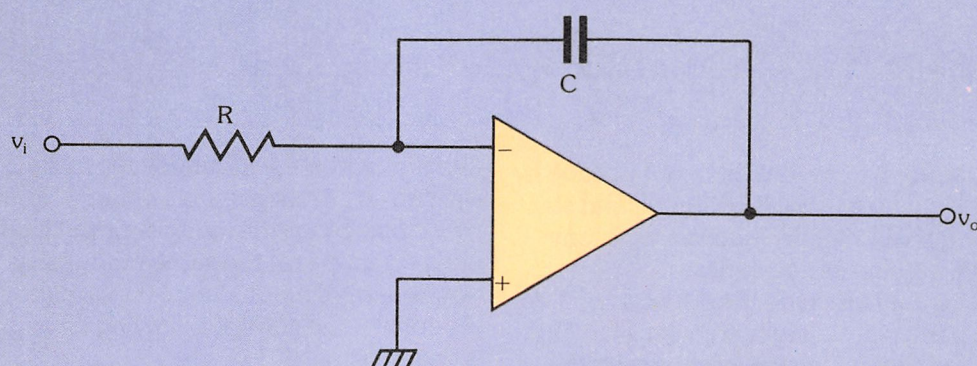
Now, the input resistance of a good op-amp is very large, so we can assume that all this current flows into the capacitor, C . This means that the charge, q , increases linearly. Thus the capacitor voltage which is equal to q/C also increases linearly. As the left-hand side of the capacitor is fixed to the virtual earth point, the right-hand side, the output voltage, must fall linearly

6



6. The integral of a sinusoidal wave is another wave lagging the first by 90° .

7



7. Constructing an integrator using an op-amp.

with time. This shows that an op-amp connected in this way can act as an integrator, while at the same time, reversing the polarity of the voltage.

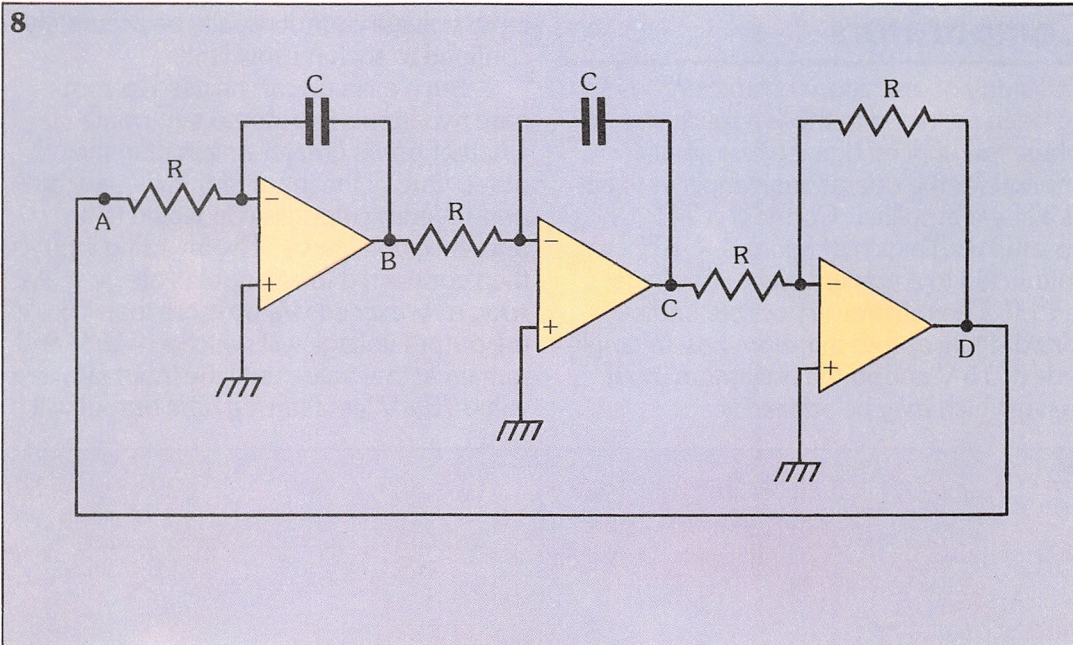
Although this has only been illustrated with a step of input voltage, the circuit will act as an integrator on any input

wave. The only approximation we have made is that a good op-amp with high gain and high input impedance is used.

Quadrature oscillator

We can use the integrator idea in constructing a sine wave generator (figure 8).

8. Constructing a sine wave generator.



This circuit consists of two integrators followed by an inverting amplifier. Let's assume that the voltage at point A is:

$$V_A = V \cos \omega t$$

then the voltage at point B will be:

$$V_B = -\frac{1}{\omega CR} (V \sin \omega t)$$

where the factor $1/CR$ is introduced by the

integrating circuit.

The voltage at point C, after going through a second integrator will be:

$$V_C = -\frac{1}{\omega CR} \left[-\frac{1}{\omega CR} (-V \cos \omega t) \right]$$

since the integral of $\sin \omega t$ is $-\cos \omega t$, then:

$$V_C = -\frac{1}{\omega^2 C^2 R^2} (V \cos \omega t)$$

Finally the voltage at point D, after passing through the inverting stage is given by:

$$V_D = -V_C = \frac{1}{\omega^2 C^2 R^2} (V \cos \omega t)$$

If the circuit is to oscillate it must satisfy the Barkhausen conditions: first, V_A and V_D must be in phase, and we can see this to be correct as they are both positive cosine waves; second, the magnitude of the loop gain V_D/V_A must be unity. This leads to:

$$\omega^2 C^2 R^2 = 1$$

and so:

$$\omega = \frac{1}{CR}$$

So, we now have a sinusoidal oscillator that operates at a frequency of:

$$\frac{1}{2\pi CR}$$

This circuit is very useful as it not only generates a sine wave at, say, point B, but a similar wave of the same amplitude, 90° out of phase is also available at point C.

Below: analogue wave generator. (Photo: Philips).



Comparators

Our study of operational amplifiers has so far been restricted to those applications where the applied signals were small enough for the operational amplifier to act as a linear amplifier. Consider a 741 op-amp with a gain of about 2×10^5 , connected to a power source of voltage ± 15 V. The maximum possible undistorted signal output therefore has an amplitude of 15 V and so the maximum input signal which may be applied is:

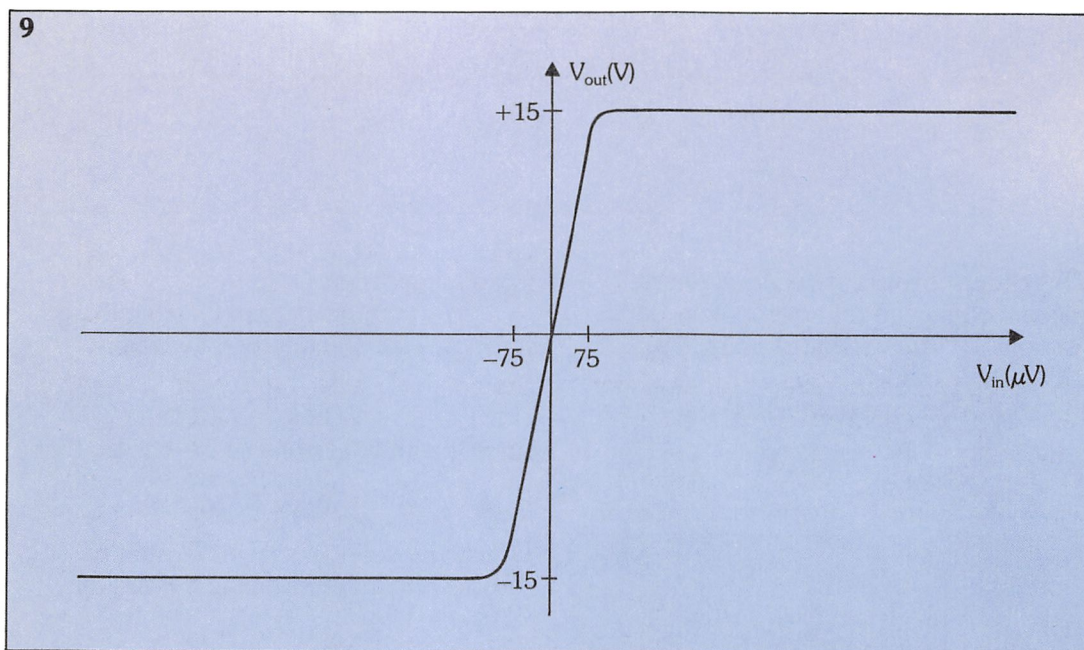
$$\frac{15}{2 \times 10^5} = 7.5 \times 10^{-5} = 75 \mu\text{V}$$

As long as the voltage between the two input terminals of the op-amp is less than $75 \mu\text{V}$, the output will be linearly related to it. If the input on, say, the non-inverting input is greater than $+75 \mu\text{V}$, the output voltage will limit at $+15$ V. In the same way, if the non-inverting input voltage is more negative than $75 \mu\text{V}$, with respect to the inverting input, then the output voltage will limit at -15 V.

This behaviour is graphically illustrated in figure 9. When the op-amp is used as a linear amplifier, the input voltage is kept small by negative feedback from the output. If this is removed then the amplifier will almost invariably limit to one or other of the power rail supply voltages, since the

input voltage cannot usually be accurately confined to such narrow limits.

Such a circuit can be used to compare two inputs in order to determine whether one is greater or less than the other. One of the input terminals, say the non-inverting input, is connected to a reference voltage V_R . The inverting input is then connected to a variable voltage v_i . As soon as v_i exceeds V_R by more than $75 \mu\text{V}$ the output voltage will switch to -15 V and remain at that value until the input falls to a value $75 \mu\text{V}$ less than V_R . The output will



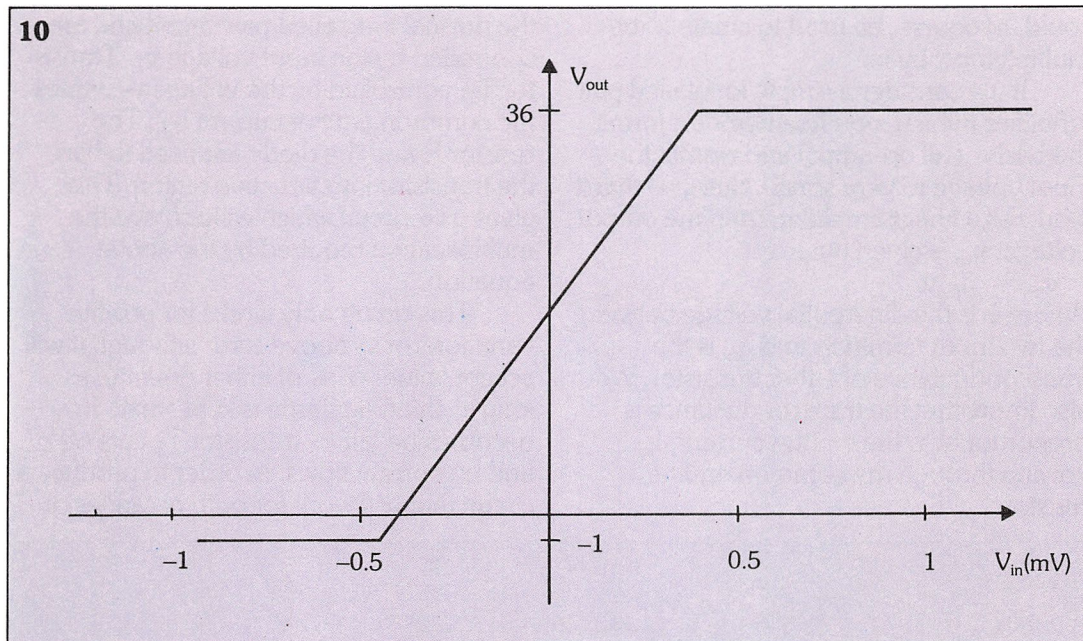
9. Transfer characteristics of a comparator switching from a low to a higher power supply rail.

then swing to $+15$ V.

Apart from a small indeterminate region around the reference voltage, the output will either be at a large positive or negative voltage. If we regarded the output as a logic circuit we could say that the output is at logic 1 ($+15$ V) when the input is less than the reference value, and at logic 0 (-15 V) when the input is greater than this value. This type of circuit is known as a **comparator**. In practice we would not rely on the power supply voltages, so it is convenient to connect two Zener diodes back-to-back across the op-amp's output, to provide a fixed limiting voltage value.

Many IC comparators exist: the 2901 comparator, for example, is compatible with TTC, DTL, MOS and CMOS logic systems, and can therefore be supplied from a single power rail at any voltage from

10. Transfer characteristic of a comparator switching around a reference voltage.



2 to 36 V. The low voltage output level is set at about -400 mV and the high level is the supply voltage.

In the range where the comparator is operating as a linear differential amplifier, it has a typical voltage gain of 10^5 , so when working from a 36 V supply it will switch from low to high for a change of input voltage of:

$$\frac{36}{10^5} = 360 \mu\text{V}$$

centred on the reference voltage. The transition region is smaller for lesser supply voltages. Figure 10 shows such a comparator's transfer characteristic.

Solution of linear algebraic equations

Earlier on we saw that an op-amp can be used to add together a number of voltages, when used with the appropriate feedback. This property can also be employed to solve a set of simultaneous algebraic equations. As an example, consider the equations:

$$x + y = 3$$

and:

$$x + 2y = 4$$

We can easily find that $x = 2$ and $y = 1$, but let's use this example to see how a circuit can be built to give the solution. First we rewrite the equations in the form:

$$y = 3 - x; \quad x = 4 - 2y$$

We may now use an op-amp to add together the two quantities 3 and $-x$ to

give $-y$ (remember the op-amp changes the sign). This is shown in black at the top of figure 11, where we have a 3 V battery and an unknown voltage $-x$ on the input terminals. The two input resistors and the feedback resistor are all equal since we need to multiply the inputs by unity. Similarly, the lower part of this diagram shows a similar op-amp arranged to give the output $-x$ from the second equation, with inputs of 4 V and $-y$. (Note the input resistor connected to the $-y$ input is $R/2$ to give the multiplying factor of 2.)

The two unknown voltages $-x$ and $-y$ exist in both of these circuits, but they are both available at the output of the opposite op-amp. If we now connect the $-x$ output to the $-x$ input and the $-y$ output to the $-y$ input (as shown in red) the circuit will be complete. When we switch the circuit on, we will find the voltage at A to be -2 V while at B it will be -1 V. Both are the negatives of the values which solve the equations. This idea is the basis behind all analogue computers.

Analogue multipliers

The type of simple op-amp setup described above can only solve linear equations, where each variable occurs on its own at no higher power than unity. Equations that involve the product of variables, $x y$, require a circuit which can multiply two variables. The same circuit

could, of course, be used to create x^2 by multiplying x by x .

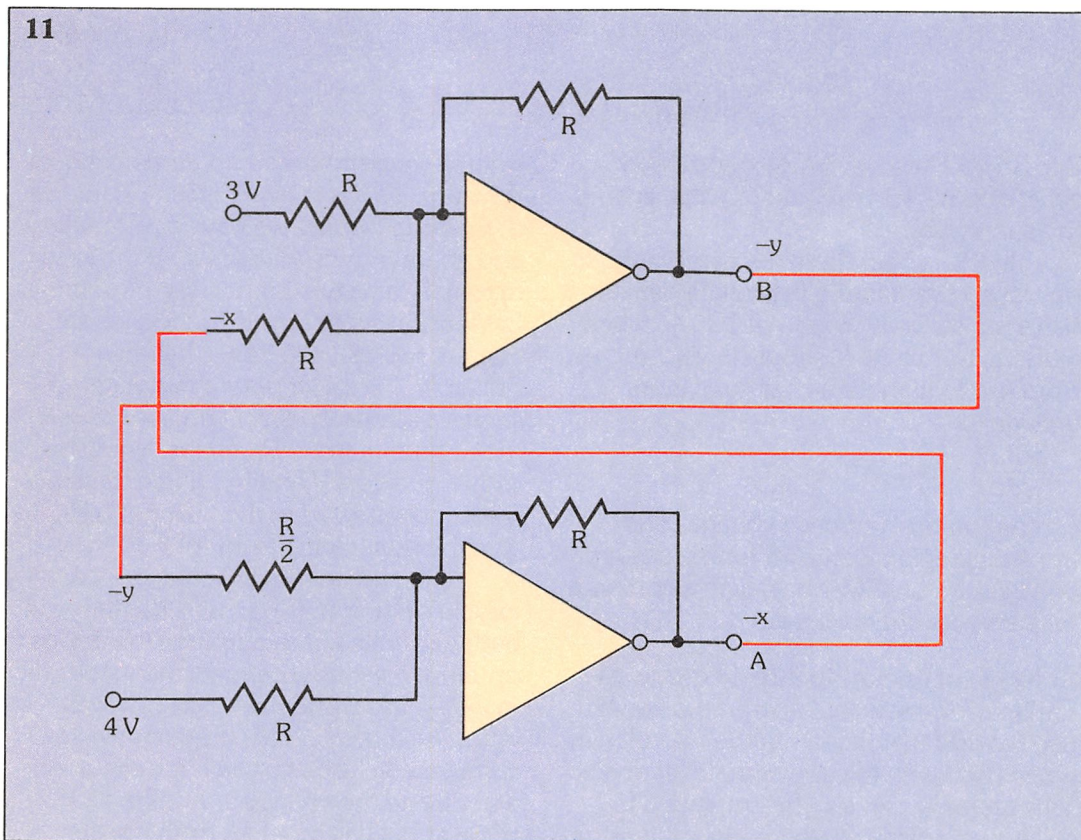
If we consider a simple longtailed pair amplifier (which, you'll remember, forms the basis of all op-amps) and restrict the input voltage to very small values, so that it works as a linear amplifier, then the output voltage, v_o , is given by:

$$v_o = g_m v_i$$

Where v_i is the differential voltage between the two input terminals and g_m is the transconductance of either transistor. We also know that the transconductance is proportional to the emitter current, I_{EE} , flowing through the common emitter resistor:

the normal longtailed pair circuit and are controlled by the input voltage v_1 . Transistor T_3 , controlled by the voltage v_2 , varies the common emitter current I_{EE} . The resistor R and the diode are used to bias the transistor into its active region. This gives us a circuit which will achieve the multiplication required by the above equation.

This circuit only works for positive variations of v_2 above zero, although it will accept values of v_1 of either polarity, as long as the magnitude is kept small. If v_2 becomes negative, transistor T_3 cuts off and no current flows. In order to produce a circuit that will work for both polarities of



11. Using an op-amp to solve linear algebraic equations.

$$g_m = \frac{e}{2kT} (I_{EE})$$

thus:

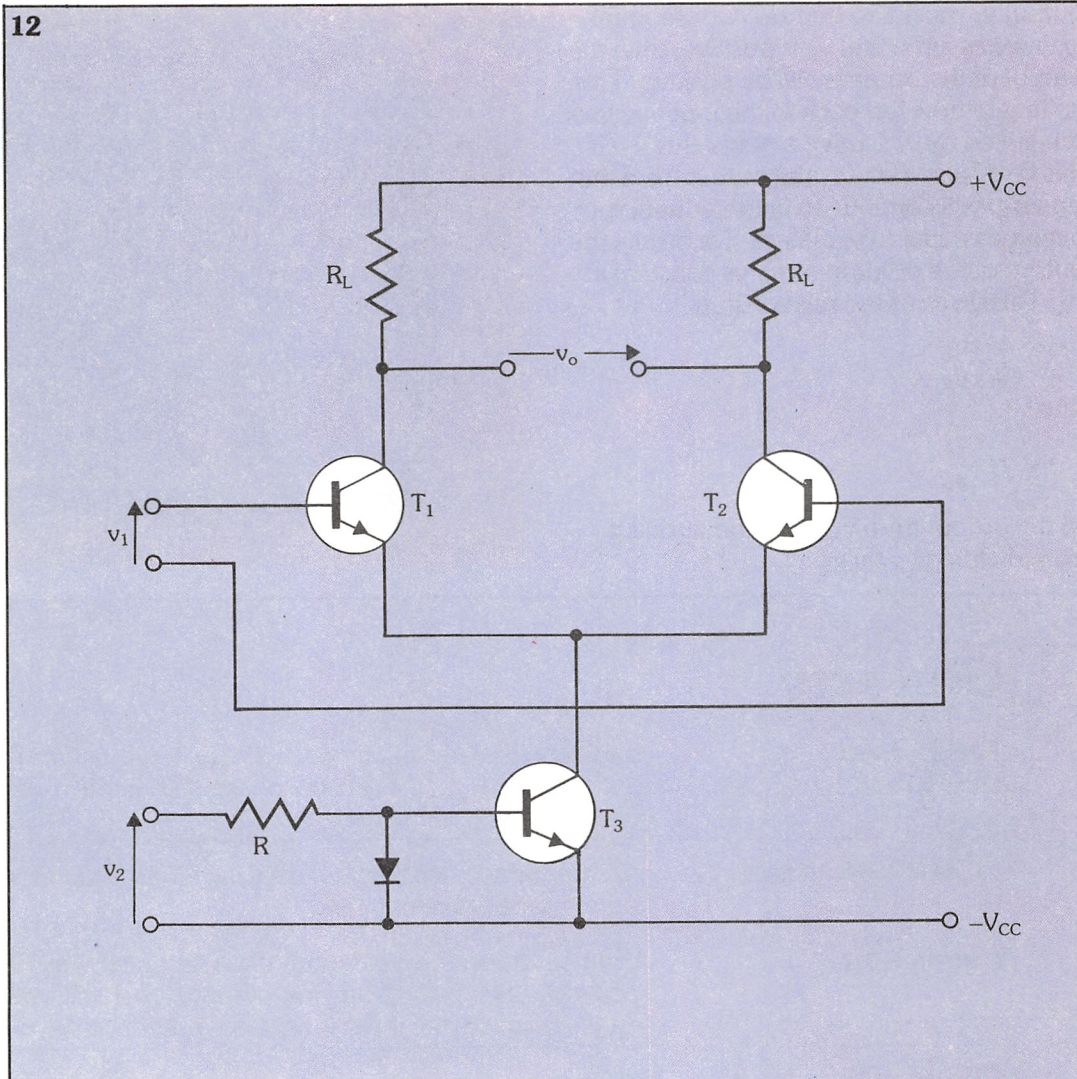
$$v_o = \frac{e}{2kT} (I_{EE} v_1)$$

If we can vary I_{EE} by a second voltage v_2 , we can make v_o depend on the product of v_1 and v_2 .

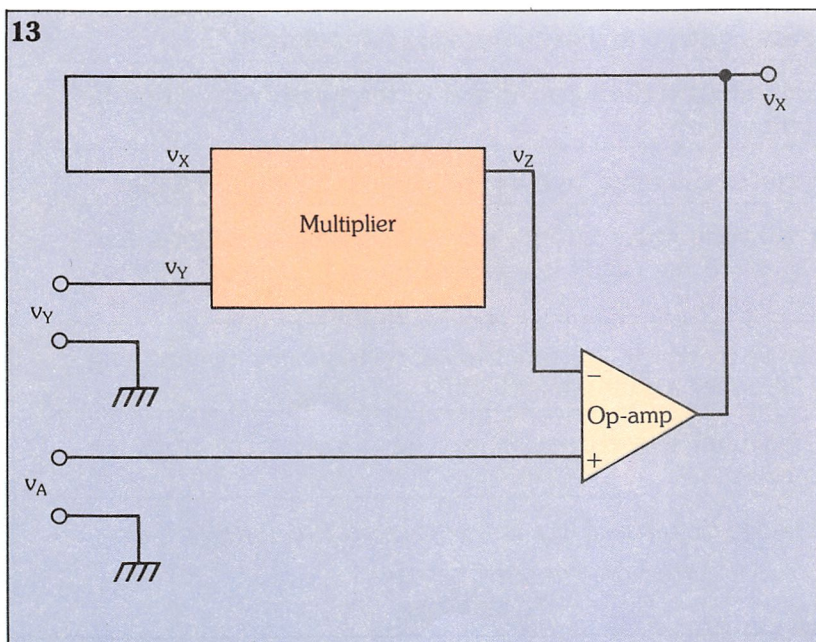
Figure 12 illustrates a circuit that can achieve all this. Transistors T_1 and T_2 form

v_1 and v_2 , we need to use two circuits similar to the one we have discussed, and then connect their bottom two transistors (T_3) together to form yet another longtailed pair. The resulting circuit is known as a **Gilbert cell**. As well as being used in analogue computation, such circuits have an important application in communications systems, where they can be used to modulate a carrier frequency by an audio or video signal.

12. Gilbert cell – used to multiply two variables together.



13. Multiplier circuit used for division.



The CA3091D is one of a number of specially designed integrated multiplying circuits. It is a monolithic circuit and operates as a four quadrant multiplier. The circuit has an accuracy of $\pm 4\%$ and operates over a very wide range of frequencies extending up to 1 MHz. It also includes circuits to improve the linearity of operation and thereby extend the voltage range over which accurate multiplication may be achieved.

Multiplier circuits can also be used to *divide* one voltage by another, if they are connected as shown in figure 13. The multiplier gives an output v_z which is the product of the two inputs v_x and v_y :

$$v_z = v_x v_y$$

The two voltages v_y and v_A are applied to the input terminals as shown. Since the voltage v_x to the multiplier is zero, the

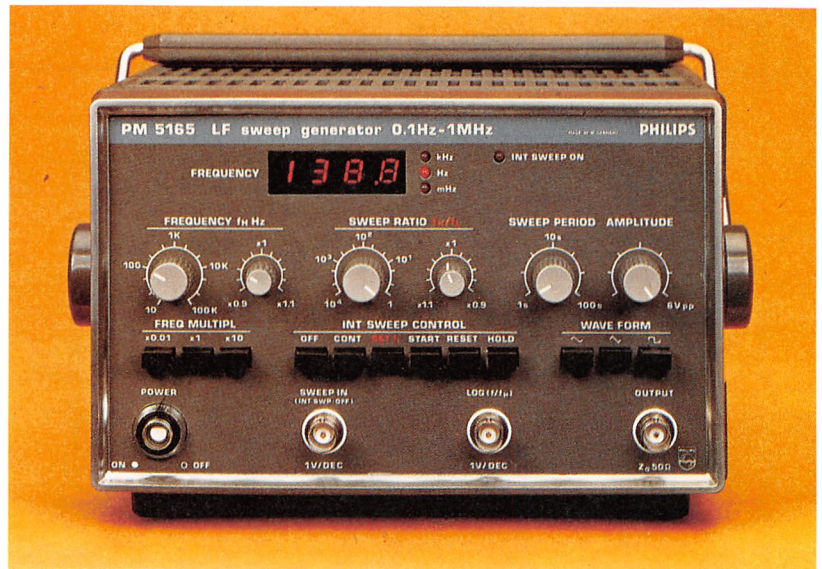
output v_z must also be zero. Consequently, if we assume that v_A is positive, then the output of the op-amp will be positive. This voltage is now fed back to the input v_x and multiplied by v_y to give a new value v_z . If this is still less than v_A , the output from the op-amp will continue to increase until v_z becomes equal to v_A . When this occurs the output of the op-amp remains constant at v_x . This gives us the relationship:

$$\begin{aligned} v_A &= v_z \\ &= v_x v_y \end{aligned}$$

thus:

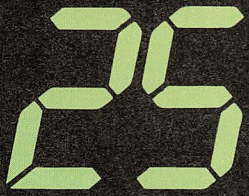
$$v_x = \frac{v_A}{v_y}$$

so the output from v_x gives the quotient from dividing v_A by v_y .



Glossary

Barkhausen conditions	conditions for oscillation which states that a network will oscillate if at some frequency, f_o , the loop phase shift is zero and the loop gain is one
Colpitt's oscillator	high frequency oscillator employing a parallel resonant circuit as its frequency selective network
comparator	circuit based on the operation of a differential amplifier. If the voltages applied to its inputs differ, the output signal will indicate which is the greater. Usually one input is held at a reference value
Gilbert cell	circuit made from two analogue multipliers that can work with both positive and negative input voltages. Has important applications in communications systems as well as analogue computation
integrator	circuit whose output is the time integral of the input wave. Used in analogue computation
multiplier	analogue circuit whose output is the product of its two input voltages
oscillator	circuit that converts direct current power. Oscillations usually start automatically when the circuit is switched on and continue until the direct voltage supply is switched off
phase shift oscillator	oscillator circuit in which the amplifier and frequency determining circuit both introduce a phase shift of 180°
quadrature oscillator	sine wave generator that consists of two integrators followed by an inverting amplifier
Wien bridge	four armed bridge circuit used as a frequency selective network



DIGITAL ELECTRONICS

Gate arrays and PLAs

Do-it-yourself ICs

Considerable technological achievements over the last few years have produced an increasing level of integration of components onto single integrated circuit chips: various reports suggest that devices are doubling in complexity every year. Electronics engineers can now make use of a greater number of components which are capable of meeting the digital requirements of both simple combinational circuits and the more complex sequential networks.

There are occasions, however, when a circuit is required which is not available in integrated circuit form, and the manufacturing quantity is not sufficient to warrant design and development of a new IC. In many cases, for example, the engineer may have produced a complex circuit which has only one function, and is usable in only one specific application.

Such a circuit may be constructed using the individual gates, flip-flops, counters, registers etc. available on individual ICs, however, it is likely that it will have considerable disadvantages, for example large size and high power consumption. Manufacturing costs may also be high due to production difficulties.

One answer to this problem is to use a type of integrated circuit, specially produced for such applications, containing a quantity of unallocated components. These integrated circuits may be thought of as being *programmable* ICs, in which the internal circuit is defined by the user. One such integrated circuit may therefore form the basis of a large number of different circuits; whichever circuit the user defines is programmed into the IC, to allocate the previously unallocated components.

We have seen throughout these *Digital Electronics* articles how a ROM device can, in fact, be used in a similar way when it acts as a code converter. Thinking back

to the calculator example in *Digital Electronics 1* (redrawn in figure 1), we know that the microprogram is stored in ROM, and that this program is arranged as a number of individual code converters. The controller stores the address of the desired instruction in the address register, and the microprogram then transfers a copy of the instruction held at this address into the instruction register – a simple code conversion. The microprogram is merely a way of holding many similar code conversions (one at each address in memory).

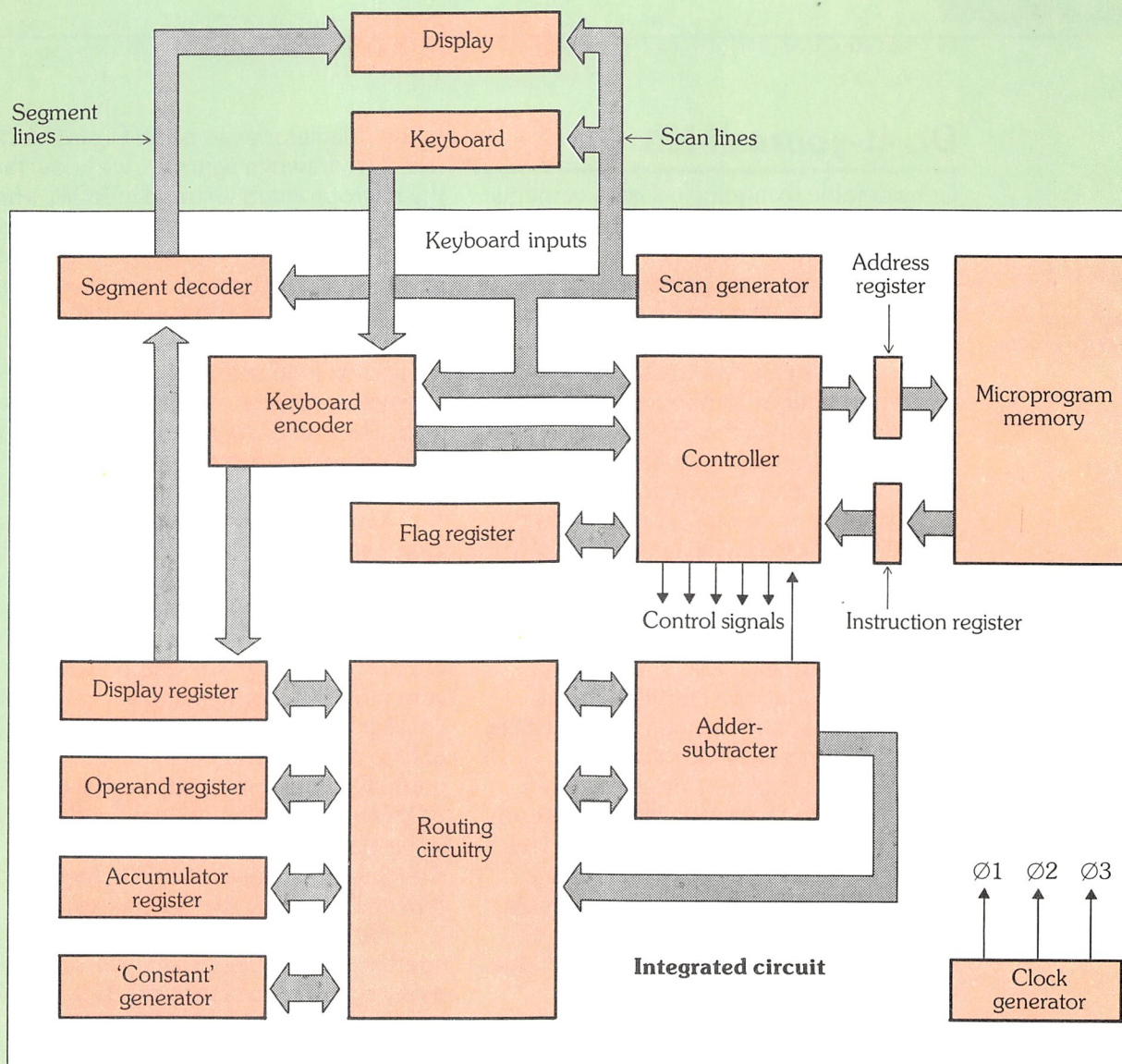
Generally, combinational circuits of logic gates are also code converters, so it is logical and correct to assume that ROM devices may be replaced by combinational circuits, and vice versa. But just think of the overall complexity, size and power consumption, of a combinational circuit which consists of a number of code converters, able to do the same job as the microprogram. Obviously, we can see that using ROM to perform the required code conversions makes things a great deal more straightforward, and a smaller, more economical circuit can be the end result.

By itself, a ROM may be used only for combinational type code conversions, but many such applications exist. It can also be used in a sequential application if some sort of controller, say, a microprocessor, is used to step it through the code converters stored. The programmable integrated circuits previously mentioned, in which the internal circuit may be user defined, form an alternative to the ROM/controller method of complex circuit configuration. This alternative may give a more cost effective solution.

There are two main categories of these programmable integrated circuits, distinguished primarily by their programming method: the **gate array**, and the **programmable logic array** (PLA; or, depending on manufacturer – PAL).

Opposite: analogue wave generators. (Photo: Philips).

1



Gate arrays are mask programmed, i.e. programmed during manufacture to a design specified by the user. Users with the necessary equipment and expertise may opt to design their own photographic masks, supplying them to the IC manufacturer during production of the gate array. Alternatively, the user simply specifies in precise detail what the integrated circuit

must do – the manufacturer then **custom designs** the gate array to suit. Programmable logic arrays, on the other hand, are field programmed by the user, by blowing preset internal fuse-links.

These two categories of programmable ICs are analogous in their programming methods to mask programmed ROMs and field programmed PROMs.

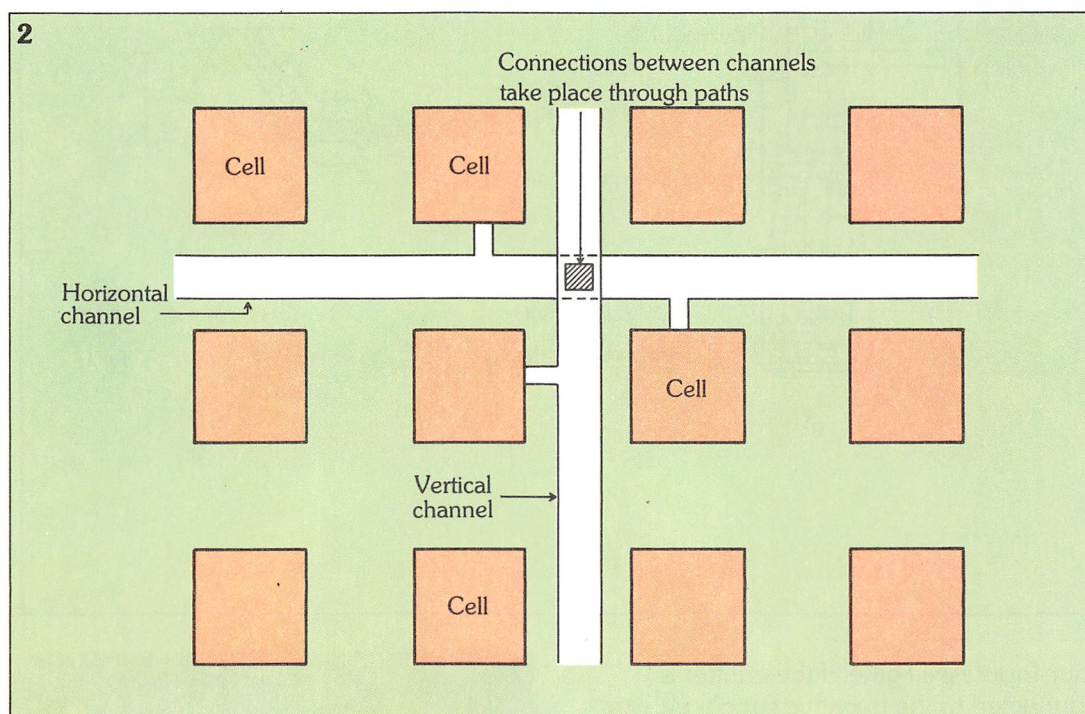
1. Simplified diagram of the major subsystems in a calculator.

Gate arrays

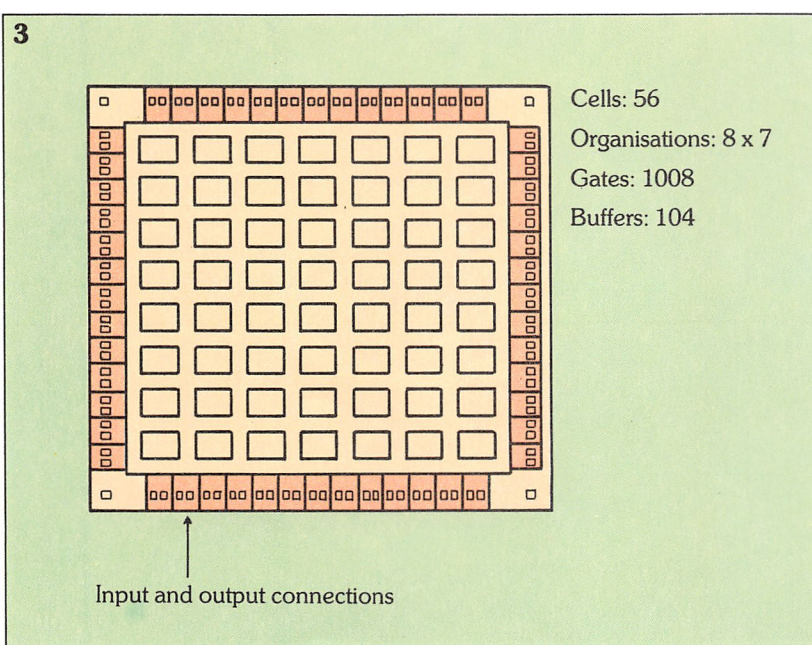
The basic structure of a gate array is shown in figure 2, and we can see that it consists of a number of individual **cells** in a lattice structure. Each cell may comprise a number of transistors or logic gates, or combinations of both. In the initial state, i.e. before programming, cells and parts of cells are isolated. The process of programming the IC consists of connecting the cells

and parts of cells, building up the required circuit. Connections are made along horizontal or vertical **channels**, formed with metallized layers in the IC. For example, a first metallization layer may form all horizontal channels, a second may form vertical channels, and a third may form power supply connections to each component or cell. Connection between channels may take place, where required, through small metallized **paths**.

2. Basic structure of a gate array showing cells in a lattice structure.



3. Cell layout of the TAT 008 gate array.

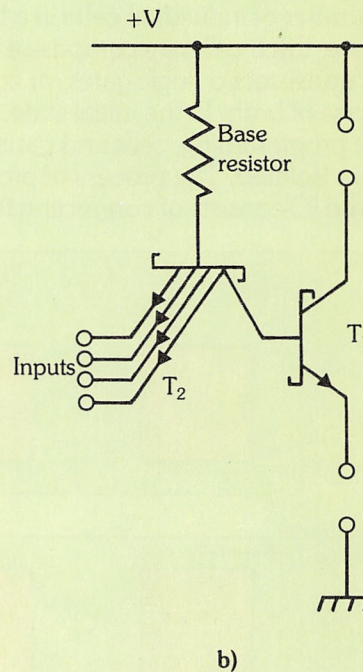
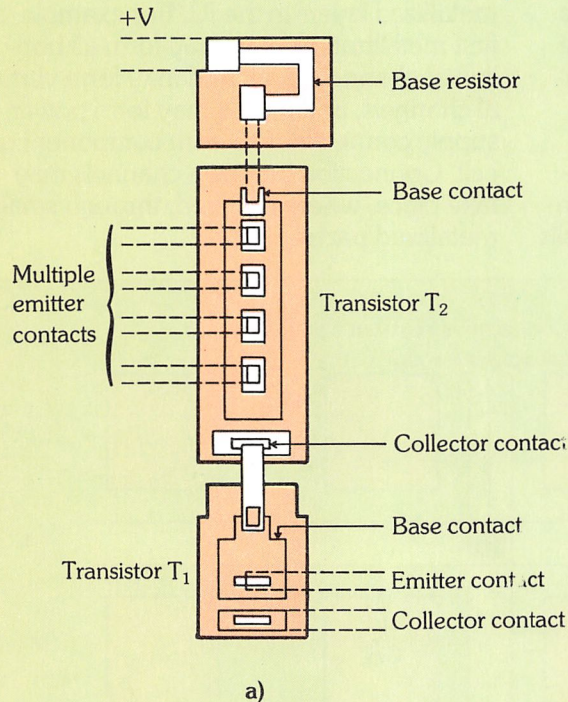


Some available gate arrays

In order to understand the manufacture and operation of gate array devices, we'll look at a couple of examples and consider their construction and use. Figure 3 shows the cell layout of a TAT 008 gate array: cells are arranged in an eight by seven array and, in fact, the IC has a total of 1008 individual gates within the 56 cells.

Figure 4a illustrates the possible layout of a basic gate used within the device, and figure 4b shows the circuit it forms. Transistor T_2 is a Schottky device (as is transistor T_1) with four inputs. The collector and emitter of transistor T_1 are initially open-circuit, allowing the user to define which is connected to the respective supply rail. If the collector is connected to the positive supply rail, the gate output is taken from the emitter and the gate forms a

4



4. (a) Basic gate within the TAT 008 gate array; (b) its circuit diagram.

Below: M²CMOS integrated circuit. (Photo: National).

four-input AND gate; if the emitter is connected to the negative supply rail, on the other hand, the gate output is taken from the collector and the gate forms a four-input NAND gate.

We have seen how all other types of logic gates – NOR, NOT and EXOR – may be constructed using only NAND and AND gates. We also know that flip-flops, registers, counters and all other sequential circuits may be built from such a range of combinational logic gates. It is therefore apparent that the basic gates in the TAT 008 may be used to construct virtually any combinational or sequential circuit that we may require.

A number of output buffer gates are also included in the gate array, which may similarly be configured in totem-pole, three-state, bidirectional or open collector mode, as required.

(continued in part 28)

